Ultrasonics 108 (2020) 106168

Contents lists available at ScienceDirect

Ultrasonics

journal homepage: www.elsevier.com/locate/ultras

Improved forward model for quantitative pulse-echo speed-of-sound imaging

Patrick Stähli, Maju Kuriakose, Martin Frenz*, Michael Jaeger

Institute of Applied Physics, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland

ARTICLE INFO

Keywords:
Ultrasound tomography
Reflection mode
Multimodal imaging
Inverse problem

ABSTRACT

Computed ultrasound tomography in echo mode (CUTE) allows determining the spatial distribution of speed-of-sound (SoS) inside tissue using handheld pulse-echo ultrasound (US). This technique is based on measuring the changing phase of beamformed echoes obtained under varying transmit (Tx) and/or receive (Rx) steering angles. The SoS is reconstructed by inverting a forward model describing how the spatial distribution of SoS is related to the spatial distribution of the echo phase shift. Thanks to the straight-ray approximation, this forward model is linear and can be inverted in real-time when implemented in a state-of-the art system. Here we demonstrate that the forward model must contain two features that were not taken into account so far: (a) the phase shift must be detected between pairs of Tx and Rx angles that are centred around a set of common mid-angles, and (b) it must account for an additional phase shift induced by the offset of the reconstructed position of echoes. In a phantom study mimicking hepatic and cancer imaging, we show that both features are required to accurately predict echo phase shift among different phantom geometries, and that substantially improved quantitative SoS images are obtained compared to the model that has been used so far. The importance of the new model is corroborated by a preliminary volunteer result.

1. Introduction

Ultrasound (US) is an integral part of today's clinical diagnostic practice since it provides us with real time display and flexible free-hand probe guidance while using non-ionizing radiation. The compact, portable and comparably inexpensive US systems make their use favourable for general practitioners, emergency units and bed-side care. On the downside, classical gray-scale B-mode US often suffers from low sensitivity and non-specific contrast, resulting in difficulties in differential diagnosis for certain disease types [1–3]. To complement B-mode images with additional structural and functional information in a single multi-modal system, much effort has been placed in developing new ultra-sound-based modalities. Apart from Doppler flow imaging, which is already state-of-the-art, recent developments include ultrasound elastography [4–9] and optoacoustic imaging [10–13].

Based on the dependence of the SoS on the tissues mechanical properties, SoS imaging is another promising modality that can help identify disease-related changes of tissue composition and structure. Breast ultrasound computed tomography (UCT) showcases the potential of SoS imaging on the example of breast cancer diagnosis [14]. In UCT, US is through-transmitted through the breast from many angles, and the spatial distribution of SoS is reconstructed based on analysing the detected US signals. In the ray approximation of ultrasound propagation, the time-of-flight (ToF) of pulsed signals is assigned to integrals of slowness (inverse of SoS) along lines connecting sender and receiver locations [15–18]. This allows a fast linear reconstruction of the SoS, e.g. via the filtered backprojection, but with the disadvantage of a low spatial resolution due to diffraction and refraction. Refraction can be accounted for in an iterative bent-ray ToF approach. Diffraction, however, requires methods that use the full signal as opposed to the ToF. Diffraction tomography based on the 1*st* order Born approximation is linear and fast, but limited to low contrast SoS variations so that its application to the breast requires a SoS prior with "reasonable low resolution" [19]. The best in vivo images were so far obtained using non-linear iterative full-wave inversion schemes, but with the disadvantage of a high numerical cost [20,21].

Whereas UCT is able to achieve high resolution and quantitative SoS images of breast cancer, the big disadvantage is that it is a through-transmission technique, which limits its use to acoustically transparent parts of the human body. To provide SoS imaging with all the flexibility of conventional handheld US and without the need for specialized equipment, SoS imaging must be based on echo US for a one-sided access. This would allow to not only image the SoS in transparent parts of the body, but also in any other part that is routinely examined using

* Corresponding author.
E-mail address: martin.frenz@iap.unibe.ch (M. Frenz).

https://doi.org/10.1016/j.ultras.2020.106168
Received 24 June 2019; Received in revised form 24 February 2020; Accepted 24 April 2020
Available online 23 May 2020

echo US, e.g. for the diagnosis of cancer other than in the breast, of fatty liver disease, or for the assessment of plaque composition inside the carotid artery.

Reflection-mode SoS imaging is feasible when taking into account a multiple-scattering process beyond the 1*st* order Born approximation. This problem can in principle be solved as in UCT in a non-linear full-wave inversion. Even though promising theoretical results were obtained using this approach in 2D digital breast phantoms [22], no clinical results have been presented so far, potentially because it requires low frequencies outside the bandwidth of conventional clinical US probes. Various alternative approaches were investigated that reconstruct the SoS based on pulse-echo signals. Techniques that estimate the average SoS between the transducer and the focal depth reported accurate SoS measurements of uniform tissue [23–25]. These methods, however, have low accuracy in the presence of SoS inhomogeneities. The crossed-beam tracking method is based on measuring the round-trip time through the intersection of two scanning beams (one for transmit, one for receive), and provides a spatial resolution on the order of 10 mm [26,27]. Recently, an approach was proposed that reconstructs the local SoS based on axial variations of the average SoS determined by optimizing the transmit and receive focusing. The technique was verified in phantoms without lateral SoS variations, where it provided an axial resolution of about 7 mm [28].

We have recently developed a reflection-mode technique, named computed ultrasound tomography in echo-mode (CUTE), which allows a real time determination of the SoS with promising spatial and contrast resolution [29,30]. CUTE is based on analysing the spatial distribution of the echo phase in beamformed (using e.g. conventional delay-and-sum algorithm) radio-frequency (rf) mode US images. A deviation of the true SoS from the value assumed for beamforming results in a mismatch between the anticipated and the actual round-trip time of US propagation (henceforth termed 'aberration delay'). A changing value of the aberration delay when detecting echoes under varying angles of ultrasound transmission and/or reception consequently results in a phase shift of these echoes, which is quantified in a spatially resolved way by e.g. Loupas type phase correlation[31]. This concept is closely related to approaches that analize the differential echo phase as function of transducer element position, to correct for aberrations caused by superficial SoS variations that act like a phase screen directly in front of the transducer aperture [32–37]. CUTE goes, however, beyond these approaches: by determining the phase shift in the beamformed images as opposed to the channel data, lateral resolution of phase shift data is achieved also away from the aperture. Based on a model of how the spatial distribution of SoS relates to the spatially resolved phase shift, the former can thus be reconstructed by solving the inverse problem. Under the straight-ray approximation, the inverse problem can be linearised via a pseudo-inverse matrix, enabling real-time SoS imaging with a spatial resolution of a couple of mm. Apart from serving as a source of diagnostic information, knowledge of the spatial distribution of SoS allows for aberration correction beyond the phase screen assumption [38,39].

The core of CUTE is the forward model that relates the SoS to echo phase shift, and this forward model is closely linked to the data acquisition scheme. Previously, we proposed an implementation of CUTE where the rf-mode images were acquired under a variety of transmit (Tx) angles, whereas the echoes were detected with a constant receive (Rx) aperture [29]. The model (henceforth termed 'old model') was based on following key assumptions:

1. Because the Rx aperture is constant, the echo phase shift is entirely determined by the changing aberration delay along the changing Tx propagation directions.
2. The echo phase shift is proportional to the difference in aberration delay along different round-trip paths.

In [29,30,38], this forward model was formulated in the frequency-domain (FD). The FD formulation, however, cannot account for the absence of phase shift data in regions of low echo intensity, which leads to artefacts in the SoS image. Therefore we proposed that quantitative imaging requires a space-domain (SD) approach [40]. Sanabria et al. consequently described a methodology for SD reconstruction from theory to implementation, and verified in a simulation study the advantage of using an SD instead of an FD approach[41], including a preliminary clinical result. A more extensive clinical pilot was presented in [42]. In spite of promising results when using both the FD and the SD approach, we realized that - for known SoS distributions - the prediction of the phase shift that is made by the old model deviates from the measured phase shift. This disagreement is shown in Fig. 1 for two exemplary scenarios, a phantom with a cylindrical inclusion inside a uniform background and a four-layer phantom. In the case of a cylindrical inclusion, the old model roughly predicts the correct phase shift magnitude, which explains that a reasonably accurate SoS contrast could be reconstructed in similar phantoms in past studies. The profile shape, however, is wrongly predicted. In case of a layered phantom, the phase shift magnitude itself is underestimated by a factor of two. In both scenarios, the old model thus fails in accurately predicting the measured phase shift. As a consequence of this mismatch between forward model and reality, using the old model for SoS reconstruction leads to inconsistent SoS results depending on phantom geometry.

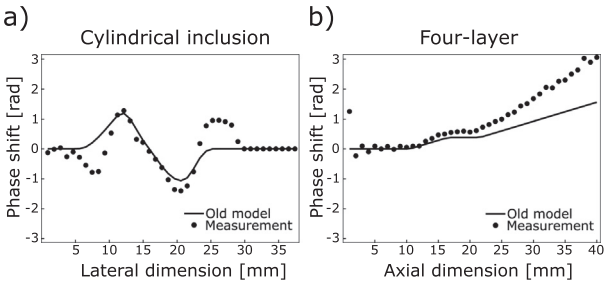

Fig. 1. Comparison of the measured phase shift to the prediction made by the old model, for a Tx angle step of $\phi_n \rightarrow \phi_{n'} = 5^\circ \rightarrow 15^\circ$, using the system described in Materials and Methods. (a) Phantom containing a cylindrical inclusion with a SoS contrast of $+30\ \mathrm{ms}^{-1}$ in a uniform background (description in Materials and Methods), lateral profile. (b) Four-Layer phantom with a SoS contrast of $135\ \mathrm{ms}^{-1}$ (description in Materials and Methods), axial profile.

Here we present a new model that solves this shortcoming by modifying the two key assumptions of the old model as follows:

1. Even if the Rx aperture is held constant while changing the Tx angle, the echo phase shift contains the influence of a virtually changing Rx angle. To avoid ambiguities, we propose to switch from a pure Tx-steering approach to simultaneously steering both, the Tx- and the Rx- angle, around a variety of common mid-angles, in an approach similar to the common mid-point gather that has been used in seismic imaging [43] as well as in US [35–37].
2. The *a priori* unknown SoS distribution leads to an offset of the reconstructed position of echoes. This offset not only depends on the aberration delay, but independently also on the values of the steering angles. As a result, the initially assumed simple proportionality between the aberration delay and phase shift is no longer valid.

In the theory section, we revise the assumptions that led to the old model and then proceed to the development of the new model. In a phantom study, we then compare the old and the new model in various different phantoms that were designed in view of two specific target clinical appliactions, i.e. layered phantoms mimicking the abdominal wall and liver (hepatic imaging) and phantoms mimicking a circular lesion inside a layered background (cancer imaging, e.g. breast). The presented results reveal that the two features of the new model are

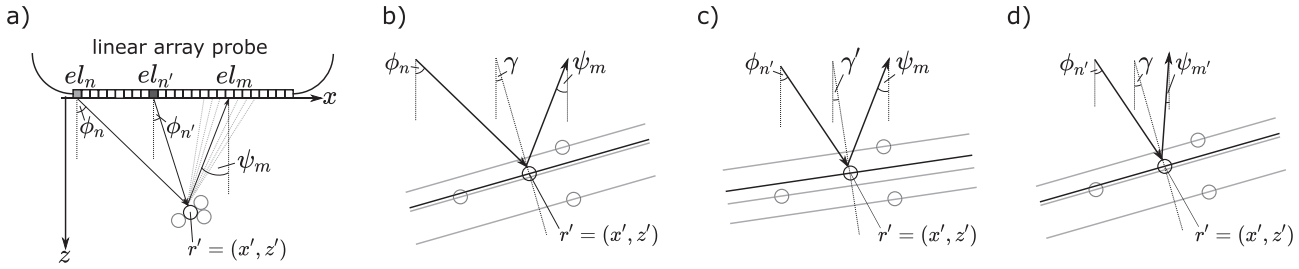

Fig. 2. (a) Sketch of a pulse-echo detection geometry. Ultrasound propagation paths are indicated, leading from different elements el_n towards a point reflector at $\mathbf{r}'$ (circles) and back to an element el_m. (b-d) Zoom-up of the region around $\mathbf{r}'$, showing the isochrones through various point reflectors when detected under different pairs of Tx and Rx angles. Only when angles are grouped around the same mid-angle (b and d) the relative position of isochrones is unchanged, thus avoiding echo decorrelation during phase tracking.

required to accurately predict the echo phase shift among different phantom geometries, and that substantially improved quantitative SoS images are obtained among all investigated phantoms.

2. Theory

The theory is developed assuming a linear array probe for 2D imaging, but can be readily adapted to curved arrays, as well as to matrix arrays for 3D imaging. Fig. 2(a) illustrates the measurement geometry. The linear array probe (parallel to coordinate x) is placed on top of the tissue (at coordinate $z = 0$), and acquires echoes using a variety of different transmit and receive settings, indexed by n and m. For the sake of visual and conceptual clarity, transmission and detection on single elements is considered in a first step in Fig. 2(a). Two different elements (indexed by el_n and $el_{n'}$) transmit an US pulse, one at a time. The two (in this case divergent) wave fronts reach a point $\mathbf{r}' = (x', z')$ inside the tissue under two different angles ϕ_n and $\phi_{n'}$. Intrinsic ultrasound reflectors located in a vicinity around $\mathbf{r}'$ lead to echoes that propagate back to the surface, where they are detected by the different sensor elements (indexed el_m) under angles ψ_m resulting in radio frequency (rf) signals $s(t, n, m)$ with the time indexed by t. Note that the important point here is not the specific type of transmit and receive beamforming: independent of the type of beamforming, ultrasound arriving at and detected from a reflector under specific angles ϕ_n and ψ_m is assumed to have propagated along paths defined by these angles. Apart from transmitting on single elements (as used in [44]), alternative types are a scanning group of elements transmitting diverging waves, plane wave transmissions having different Tx angles [29,30], or weakly focused or collimated scanning beams (as in a classical line-by-line scan). Similarly, alternatives to using single elements for reception are e.g. defocused groups of elements, steered collimated or focused groups of elements, steered receiving plane waves (or frequency-domain filtering of rx angles, as used in this study, see later on), etc.

2.1. The old model

Given a combination of transmit and receive setting (n, m), the complex (analytic) rf-signal (crf-signal) generated by a single point reflector located at $\mathbf{r}'$ can be modelled as a product of a complex exponential carrier modulated with a complex-valued envelope G:

$$s(t, n, m) = G(t - t_0(\mathbf{r}', n, m)) \exp[2\pi i f_0 (t - t_0(\mathbf{r}', n, m))] \quad (1)$$

where f_0 is the centre frequency, t_0 is the actual round-trip time of the echo, and the function G describes the complex envelope of the signal. We thereby assume that G is determined by the spatio-temporal impulse response of the US system alone (i.e. not influenced by the tissue's acoustic properties) and that any dependence on n or m has been removed by calibration. The old model was based on the following rationale: the delay-and-sum (DAS) beamformed crf-amplitude in a point $\mathbf{r}$ at the true position of the point reflector $\mathbf{r}'$ is:

$$\begin{aligned} &u(\mathbf{r} = \mathbf{r}', n) = \\ &\textstyle\sum_m [s(\hat{t}_0(\mathbf{r}, n, m), n, m)] = \sum_m [s(\hat{t}_0(\mathbf{r}', n, m), n, m)] = \\ &\textstyle\sum_m \{G((\hat{t}_0 - t_0)(\mathbf{r}', n, m)) exp[2\pi i f_0 (\hat{t}_0 - t_0)(\mathbf{r}', n, m)]\} \end{aligned} \quad (2)$$

where $\hat{t}_0$ is the anticipated round-trip time for point $\mathbf{r}$, calculated based on an *a priori* SoS $\hat{c}$. Deviations of the true SoS $c(\mathbf{r})$ from the anticipated value $\hat{c}(\mathbf{r})$ lead to a deviation of the actual round-trip time from the anticipated value. This deviation, termed aberration delay, consists of two parts: one is the delay $\tau_{tx}(\mathbf{r}', n)$ of the n-th transmitted wave front when propagating towards $\mathbf{r}'$, and the other one is the delay $\tau_{rx}(\mathbf{r}', m)$ of echoes propagating from $\mathbf{r}'$ to the m-th receiving element, so that $t_0(\mathbf{r}', n, m) = \hat{t}_0(\mathbf{r}', n, m) + \tau_{tx}(\mathbf{r}', n) + \tau_{rx}(\mathbf{r}', m)$. With these aberration delays, Eq. 2 becomes ($\mathbf{r}'$ is omitted for notational simplicity):

$$\begin{aligned} &u(n) = \\ &\textstyle\sum_m \{G(-\tau_{tx}(n) - \tau_{rx}(m)) exp[2\pi i f_0 (-\tau_{tx}(n) - \tau_{rx}(m))]\} \end{aligned} \quad (3)$$

Phase shift tracking is based on the following assumption: G varies 'slowly' compared to the oscillations of the exponential carrier, either implicitly due to the bandlimited frequency response of typical clinical US probes, or explicitly by bandpass filtering. For a sufficiently small difference between Tx angles ϕ_n and $\phi_{n'}$, the change in τ_{tx} will be small compared to the temporal variation of G, so that the value of G can be assumed constant and τ_{tx} can be replaced by its average $\overline{\tau_{tx}}$ in the envelope. Eq. 3 is then simplified to:

$$\begin{aligned} &u(n) \simeq \exp(-2\pi i f_0 \tau_{tx}(n)) \\ &\textstyle\sum_m [G(-\overline{\tau_{tx}} - \tau_{rx}(m)) \cdot \exp(-2\pi i f_0 \tau_{rx}(m))] \end{aligned} \quad (4)$$

On the right-hand side of Eq. 4, only the complex pre-factor depends on n. The first processing step of CUTE when using the old model is the experimental determination of the change in this pre-factor as a function of n. This is achieved by determining a map of local echo phase shift $\Delta\Theta(\mathbf{r}, n, n')$, defined as the phase angle of the (locally averaged) point-wise Hermitian product between the crf-mode images obtained with Tx settings n and n'. At point $\mathbf{r} = \mathbf{r}'$:

$$\begin{aligned} &\Delta\Theta(\mathbf{r}', n, n'): = \\ &\arg\left[\int_{x'-0.5\Delta x}^{x'+0.5\Delta x} \int_{z'-0.5\Delta z}^{z'+0.5\Delta z} d\mathbf{r}\{u(\mathbf{r}, n) \cdot u^{\star}(\mathbf{r}, n')\}\right] \end{aligned} \quad (5)$$

The averaging over a limited area around $\mathbf{r}'$ ('tracking kernel'), with size Δx by Δz, improves the robustness of the phase determination. The size of the tracking kernel defines the trade-off between spatial and contrast resolution of CUTE. The old model is based on the assumption that the experimentally measured phase shift of an echo is determined by the phase shift at the true location $\mathbf{r}'$ of the echo-generating reflector. Under this assumption and provided that the change in τ_{tx}, $\Delta\tau_{tx}$, is smaller than half the oscillation period $1/f_0$ (to avoid phase aliasing), the measured echo phase shift is – according to Eq. 4 – related to the Tx aberration delay $\tau_{tx}(\mathbf{r}', n)$, as:

$$\Delta\Theta(\mathbf{r}', n, n') \simeq 2\pi f_0 \Delta\tau_{tx}(\mathbf{r}', n, n') = 2\pi f_0 \{\tau_{tx}(\mathbf{r}', n') - \tau_{tx}(\mathbf{r}', n)\} \quad (6)$$

Note that a $\Delta\tau_{tx}$ that fulfils the aliasing condition automatically also fulfils the condition for Eq. 4. A simple formula for the prediction of the aliasing limit given the angle step between ϕ_n and ϕ_n' does not exist, as it depends on the aberration delays and thus on the *a priori* unknown spatial distribution of the SoS. In practice, one can base such a prediction on simulating the τ_{tx} for an expected variety of possible SoS distributions, or directly on observing the presence/absence of aliasing in the experimental phase shift data.

To complete the forward model, the aberration delays need to be related to the spatial distribution of SoS. Similar to UCT, including US diffraction and refraction in the forward model is in principle feasible, but would result in a time-consuming iterative reconstruction. In view of real time SoS imaging, we thus adhere to the straight-ray approximation and relate $\tau_{tx}(\mathbf{r}', n)$ to line integrals of slowness deviation $\Delta\sigma(\mathbf{r}, n)$, along propagation angles ϕ_n – determined by the transmit setting n – from the transducer surface to point $\mathbf{r}'$:

$$\tau_{tx}\left(\mathbf{r}', n\right) = \int_{\phi_n}^{\mathbf{r}'} dl \left\{\frac{1}{c(\mathbf{r})} - \frac{1}{\hat{c}(\mathbf{r})}\right\} \equiv \int_{\phi_n}^{\mathbf{r}'} dl \Delta\sigma(\mathbf{r}) \quad (7)$$

Given the abundance of point reflectors in tissue, the function $\Delta\Theta(\mathbf{r}', n, n')$ can be measured not only at one point but (ideally) throughout the imaging plane. The old model is based on the assumption that the phase shift measured from any echo corresponding to a point reflector is independent from the absence/presence of nearby reflectors. By inverting the forward model consisting of Eq. 6 and Eq. 7, an estimated spatial distribution of SoS $c'(\mathbf{r})$ can then be reconstructed from $\Delta\Theta$.

2.2. New model

As mentioned in the Introduction, the new model involves two fundamental changes to the CUTE methodology: in a first step, the common mid-angle (CMA) approach is motivated. In a second step, we take into account the position offset of the reconstructed echoes.

2.2.1. Common mid-angle tracking

A first important prerequisite for the old model expressed in Eq. 6 was the assumption that G did not depend on n nor m, and thus the signal depended on n and m only via the aberration delays. In reality, however, G itself can depend on n and m: (a) lateral gradients of the distribution of SoS result in wave front aberrations that distort the shape of G in a spatially non-uniform way. We continue assuming that SoS variations are moderate enough so that this part is negligible. (b) The spatial distribution of the tissue's reflectivity function leads to coherently interfering echoes. The influence of the tissue's reflectivity function can by no means be neglected, as it is precisely the echoes generated by this function that CUTE is based on. The effect this can have is illustrated for two extreme (but commonly found) cases:

- Specular reflector: the echo from a reflector intersecting with $\mathbf{r}'$ propagates back to the probe along a direction ψ – determined by the mirror law – that varies opposite to a varying ϕ_n.
- Uniform diffuse scattering: The echoes from a dense (below spatial resolution cell) and random distribution of 'point' reflectors interfere at the aperture. The signal detected at the aperture decorrelates with changing n, because the relative round-trip times of the different echoes change.

In both cases, and in a more general scenario, the spatial distribution of the tissue's reflectivity function leads to a dependence of G itself on n and m, and the relation between the aberration delay and the measured echo phase shift becomes ambiguous.

To avoid such ambiguities, we introduce a fundamental change to the CUTE methodology: common mid-angle (CMA) tracking. In this technique, crf-mode images are reconstructed for pairs of Tx angles ϕ_n and Rx angles ψ_m that have the same mid-angle. Similar to the common mid-point approach that was previously used in aberration correction [38], the CMA approach makes use of the following fact: signals obtained under angle pairs $(\phi_n|\psi_m)$ grouped around the same mid-angle $\gamma = 0.5(\phi_n + \psi_m)$ are well correlated, independent of the spatial distribution of US reflectors. For illustration, Fig. 2(b) shows a scale-up of a small region around point $\mathbf{r}'$, containing a number of point reflectors. Each reflector is the pivot of an isochrone, a line that connects all points that would lead – for a specific pair $(\phi_n|\psi_m)$ – to the detection of an echo at an identical time. We assume that the curvature of the isochrones (which may result from the beamforming and/or from wave front distortions due to a non-uniform SoS) is negligible over the size of the observed region, so that the isochrones can be approximated as straight lines. When changing ϕ_n to $\phi_{n'}$ without changing ψ_m (Fig. 2(c)), the distances between the isochrones change, leading to a changing interference of the reflected echoes and thus to decorrelation of the signal that is detected from direction ψ_m. When changing ψ_m in opposite direction to $\psi_{m'}$ (Fig. 2(d)) such that the mid-angle in the pairs $(\phi_n|\psi_m)$ and $(\phi_{n'}|\psi_{m'})$ is fixed (so that $\gamma = 0.5(\phi_n + \psi_m) = \gamma' = 0.5(\phi_{n'} + \psi_{m'})$), then the relative distances between isochrones are not altered, and thus the signal is well correlated. Note that the isochrones do not need to be constructed as part of the theoretical derivations, but are merely used as a didactic concept to illustrate why CMA tracking leads to correlated echoes independent of tissue reflectivity. The CMA approach can also be understood along the perspective of the k-space representation of the pulse-echo system response [45]: the system response corresponding to an angle pair $(\phi|\psi)$ detects a line in the 2D spatial Fourier transform of (an area of) the tissue's reflectivity function, along k-vectors pointing into the direction given by the mid-angle $0.5(\phi + \psi)$. Crf images (or an area thereof) acquired with angle pairs resulting in the same mid-angle thus sample the same line of the k-space representation of the reflectivity function, resulting in correlated echoes. Angle pairs corresponding to different mid-angles sample different lines in k-space resulting in uncorrelated echoes (unless the reflectivity function itself is correlated in k-space such as in the case of a single point reflector).

To employ the CMA approach, we reconstruct crf-mode images $u(\mathbf{r}, n, g)$ where n indexes the Tx setting resulting in a Tx angle ϕ_n as before, and the new index g indexes the mid-angle γ_g out of a set of mid-angles. Both together define the receive angle $\psi_{n,g}$: $=\psi_{m(n,g)} = 2\gamma_g - \phi_n$. This way of indexing the Rx angle takes into account that – for a given mid-angle – the Rx angle cannot be freely chosen, but is determined by the Tx angle. Accordingly, the phase correlation is adapted from Eq. 5 to:

$$\Delta\Theta(\mathbf{r}', n, n', g) = \arg\left[\int_{x-0.5\Delta x}^{x+0.5\Delta x}\int_{z-0.5\Delta z}^{z+0.5\Delta z} d\mathbf{r}(u(\mathbf{r}, n, g)\cdot u^{\star}(\mathbf{r}, n', g))\right] \quad (8)$$

CUTE is based on evaluating echo phase shifts in crf-mode images as opposed to channel data because this better allows to spatially resolve the influence of the SoS on the echo phase shift of different echoes. In the old model, the lateral resolution of the crf-mode images was achieved via Rx focusing (weak synthetic Tx focusing was added mainly to reduce clutter noise). In the CMA approach, when reconstructing images with sharp ϕ_n and ψ_m, no lateral resolution would be obtained. Therefore, to achieve lateral resolution, a non-zero Rx angular aperture must be used together with a non-zero Tx angular aperture for combined Tx- and Rx beamforming. In comparison to the old model where the Rx angular aperture could in principle make use of the full transducer aperture, it must be chosen substantially smaller in the CMA approach so as to resolve different Rx steering angles. One has to keep in mind that reducing the Rx angular aperture goes hand-in-hand with a reduced lateral resolution of the crf-mode images, limiting in turn the

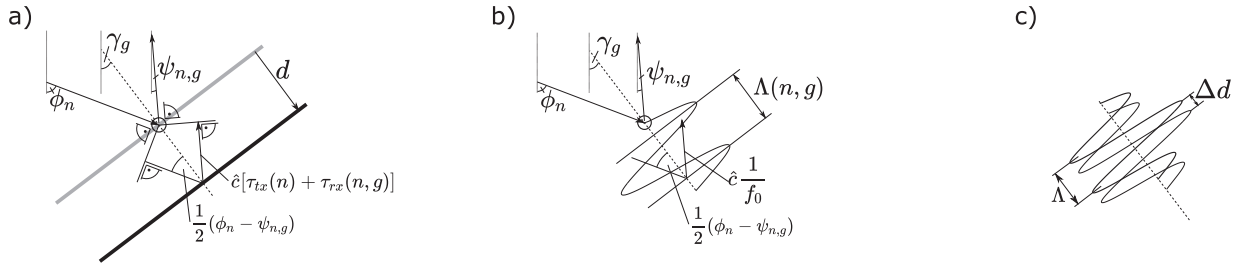

Fig. 3. (a) Geometrical sketch showing the relation between a Tx/Rx angle pair ($\phi_n|\psi_{n,g}$), the aberration delays ($\tau_{tx}(n)$, $\tau_{rx}(n, g)$), and the spatial offset d of an isochrone. (b) Similar sketch showing the relation between the centre frequency f_0 and the echo's spatial oscillation period $\Lambda(n, g)$. (c) The measured phase shift is determined by the ratio between the shift $\Delta d(n, n', g)$ of the offset of the echo and Λ.

lateral resolution of the final SoS image. We consider this an acceptable drawback of CMA tracking, compared to the big advantage of enabling improved SoS accuracy.

2.2.2. Echo position offset model

A second important assumption of the old model was the evaluation of the echo phase shift at the true location of the ultrasound reflector ($\mathbf{r} = \mathbf{r}'$). It is, however, not the phase shift at the true location of the reflector that determines the value of the measured phase shift, but the phase shift due to the spatial shift of the reconstructed position of the echo. As a second fundamental modification to the CUTE methodology, this is taken into account by the new model.

To derive the new model of how aberration delays relate to the echo phase shift, we again use the concept of the isochrones (Fig. 3). We assume a reflector (point or plane) that is detected with a specific angle pair ($\phi_n|\psi_{m,n}$). The total aberration delay (in Fig. 3(a) assumed larger than zero) leads to an offset of the reconstructed position of the echo, away from the true location of the reflector, to the isochrone for which the anticipated round-trip time agrees with the actual round-trip time. The offset d between the location of the reflector and the isochrone is – according to the geometric considerations depicted in Fig. 3(a) – given by:

$$d\left(n, g\right) = \frac{\hat{c}\,[\tau_{tx}(n) + \tau_{rx}(n, g)]}{2\cos\left[\frac{1}{2}(\phi_n - \psi_{n,g})\right]} \tag{9}$$

The phase shift is determined by the spatial shift of the offset. Formally, when changing ϕ_n by $\Delta\phi$ to $\phi_{n'}$ (so that $\psi_{n,g}$ changes by $-\Delta\phi$ to $\psi_{n',g}$), d changes accordingly by:

$$\Delta d = \frac{\hat{c}\,[\tau_{tx}(n') + \tau_{rx}(n', g)]}{2\cos\left[\frac{1}{2}(\phi_{n'} - \psi_{n',g})\right]} - \frac{\hat{c}\,[\tau_{tx}(n) + \tau_{rx}(n, g)]}{2\cos\left[\frac{1}{2}(\phi_n - \psi_{n,g})\right]} \tag{10}$$

To derive the relation between Δd and the phase shift, the spatial oscillation period of the reconstructed echo is required. This can follow the same argumentation as for d (see Fig. 3(b)): the spatial period Λ is given by the distance between two isochrones that are separated by one temporal period $1/f_0$:

$$\Lambda\left(n, g\right) = \frac{\hat{c}/f_0}{2\cos\left[\frac{1}{2}(\phi_n - \psi_{n,g})\right]} \tag{11}$$

The fact that Λ varies with varying angles implies that (i) the phase shift varies along the envelope of an echo and (ii) the interference of overlapping echoes partially decorrelates, both resulting in phase noise. To avoid this potential source of errors (and to simplify maths), we assume that the signals are bandpass filtered in a way that the spatial period Λ does not change between n and n' (for one way how this can be achieved, see Materials and Methods) and takes, for example, the value:

$$\Lambda = \frac{\hat{c}/f_0}{2} \tag{12}$$

With this convention, the new model for the phase shift is:

$$\Delta\Theta(\mathbf{r}', n, n', g) \simeq \frac{2\pi\Delta d}{\Lambda} =$$
$$2\pi f_0 \left\{ \frac{\tau_{tx}(n') + \tau_{rx}(n', g)}{\cos\left[\frac{1}{2}(\phi_{n'} - \psi_{n',g})\right]} - \frac{\tau_{tx}(n) + \tau_{rx}(n, g)}{\cos\left[\frac{1}{2}(\phi_n - \psi_{n,g})\right]} \right\} \tag{13}$$

In comparison to the old model, the new model (Eq. 13) not only includes CMA tracking, but also an inverse-cosine law that accounts for the influence of the angles on the echo position offset. Note that the inverse-cosine law in Eq. 13 does not depend on the assumption made in Eq. 12: even if Λ is allowed to change together with d, a very similar expression is obtained where Λ is replaced by the average value of the Λ for the different angle pairs (maths not shown to avoid distraction). The only disadvantage of allowing Λ to vary could be, as said, an increase in phase noise.

Similar to the old model, the equality in Eq. 13 only holds if the angle step is chosen so as to avoid phase aliasing. Note that a prediction of the aliasing limit can no more be based on the aberration delays directly in the way this was suggested for the old model, but must take into account the effect of the inverse-cosine law. It is thus ideally based on the complete right-hand side of Eq. 13.

3. Materials and methods

3.1. Space domain forward and inverse model

The theory section has been formulated with the intention to be general without restriction to specific ways of Tx/Rx beamforming (plane wave, diverging wave, focused line-by-line, etc.). For the experimental study, however, we narrow our focus down to a plane wave Tx approach. In comparison to a diverging wave approach (as, for example, used in [44]), this has the advantage of an improved SNR and that - for a specific Tx setting n - all points inside the tissue (apart from the Tx shadows explained later on) are insonified with an identical Tx angle. In the theory section, we did also not limit ourselves to specific Tx/Rx angle pairs. We only mentioned that the Tx/Rx angle steps need to be sufficiently small so as to avoid phase aliasing. This criterion depends on the specific SoS contrast, and for the phantoms used in our study, it was fulfilled with a 2° angle step (see further below). One could base the SoS reconstruction on phase shift maps acquired with the same angle resolution, but – at the same time – tracking over a maximum possible angle range is desired as this reduces ill-posedness of the SoS reconstruction and thus improves robustness and spatial resolution of the SoS image. The SoS reconstruction would thus have to be based on a large number of phase shift maps resulting in an unnecessarily high computational cost. To reduce the number of phase shift maps and thus the computational cost, we accumulate the phase shift to a coarser angle resolution. Because the final Tx/Rx/mid angles used for SoS reconstruction determine the whole processing chain, we start by

outlining these parameters before presenting the different data acquisition and processing steps.

To further simplify the numerics, we choose an equidistantly spaced Tx angle set $\{\phi_n\}$. For CMA tracking, the equidistant spacing of ϕ_n allows choosing a set $\{\gamma_g\}$ of common mid-angles in a way that – for each ϕ_n – the set of Rx angles $\{\psi_{n,g}\}$ can be made identical to the set $\{\phi_n\}$ (this would not be possible with non-equidistant spacing). For clarity, we say that we choose ϕ_n and $\psi_{n,g}$ so that they share the same angle set $\{\Phi_l\}$, i.e. $\phi_n \in \{\Phi_l\}$ and $\psi_{n,g} \in \{\Phi_l\}$. This choice allows to make use of data redundancies as explained further below, but it also simplifies the indexing: given that the Rx angle sets $\psi_{n,g}$ are identical, independent of ϕ_n, it is now practical to index the Rx angle as ψ_m (instead of $\psi_{n,g}$), and corresponding mid-angle as $\gamma_{n,m}$. The reader can convince her/himself that the sets of mid-angles $\{\gamma_{n,m} = 0.5(\phi_n + \psi_m)\}_n$ corresponding to the different ϕ_n are intersecting subsets of one set $\{\gamma_g\}$ of common mid-angles.

For this study, $\{\Phi_l\}$ is $\{-25°, -15°, -5°, 5°, 15°, 25°\}$ for $l = [1 \ldots 6]$. Thereby, the 10° step was empirically chosen as a good compromise between computational cost and SoS contrast resolution, and the ±25° are limited by grating lobes that become dominant at larger angles (determined by the array probe's element pitch in relation to center frequency). As the goal of the experimental part is to compare the old and the new model, both Tx-only tracking (old model) and CMA tracking (new model) are used in this study. For Tx-only tracking, phase shift is accumulated between successive Tx angles ϕ_n and ϕ_{n+1}, leading to phase shift maps $\Delta\Theta(\mathbf{r}', n)$ for $n = [1 \ldots 5]$. For CMA tracking, Fig. 4 (a) summarizes the resulting combinations of ϕ_n, ψ_m and $\gamma_{n,m}$. As one can see, the $\gamma_{n,m}$ are all part of a set $\{\gamma_g\}$ equal to $\{-25°: 5°: 25°\}$. Phase shift is accumulated between angle pairs with common mid-angles $\gamma_{n,m}$, indicated in Fig. 4 (a) by arrows pointing along fields having the same gray value. Since only one pair $(\phi_n|\psi_m)$ exists for mid angles $\gamma_{n,m} = \pm 25°$, no tracking is performed for these $\gamma_{n,m}$. The tracking procedure results in phase shift maps $\Delta\Theta(\mathbf{r}', n, m)$, for $n, m \in [1 .. 5]$, between angle pairs $(\phi_n|\psi_m)$ and $(\phi_{n+1}|\psi_{m-1})$. Fig. 4 (b) exemplarily shows the phase shift maps from CMA tracking of a cylindrical inclusion phantom (see description of phantom later on). These maps are intuitively arranged in a 2D matrix, according to the different angle combinations shown in Fig. 4 (a). As mentioned above, choosing identical sets of Tx and Rx angles for CMA tracking allows making use of data redundancy: in an ideal case, the crf-mode images resulting from interchangeable $(\phi_n|\psi_m)$ angle pairs $(\Phi_l|\Phi_{l'})$ and $(\Phi_{l'}|\Phi_l)$ are identical by time reversal symmetry. As a consequence, the phase shift maps between $(\Phi_l|\Phi_{l'})$ and $(\Phi_{l+1}|\Phi_{l'-1})$ and the ones between $(\Phi_l|\Phi_{l'})$ and $(\Phi_{l'-1}|\Phi_{l+1})$ provide redundant data. This is illustrated by the experimental data in Fig. 4 (b), where the phase shift maps corresponding to interchanged angle pairs, i.e. the maps located symmetrically about the diagonal, are indeed very similar apart from a sign change due to the change in tracking direction. The observed differences in the spatial distribution of the phase shift between redundant maps hint at a residual asymmetry, which is expected in practice for slight differences in the implementation of Tx and Rx beamforming (as in our study, see later). Residual asymmetry is best observed in the phase shift maps corresponding to identical angle combinations, i.e. the maps in the diagonal of Fig. 4 (b): with perfect symmetry, these maps would ideally be zero, and - conversely - they can serve as a measure of symmetry of Tx/Rx beamforming. Due to the redundancy described above, only either the maps below (framed in Fig. 4 (b) by a red border) or above the diagonal (framed in Fig. 4 (b) by a blue border) are in principle needed for SoS reconstruction. To reduce the observed deviations from symmetry, however, it is advantageous to generate all possible maps and then average the redundant maps after sign inversion (as it is done in this study).

With a specific set of beamforming parameters (e.g. Tx angle and angular aperture, Rx angle and angular aperture) only part of the imaging plane can be detected due to the limited aperture size of the probe. This leads to areas of missing data ('shadows') towards the lateral edges of the image. In Tx-only tracking (old model) where the full Rx aperture is used, the shadows are determined by the Tx angle alone. In CMA tracking, the combined Tx- and Rx-shadows have to be taken into account. In both cases, when tracking between angles, valid phase shift data is only available in the area where echoes are available before and after the angle step. For CMA tracking (where both Tx and Rx are steered), the combined Tx/Rx shadows are indicated in 4 (b) as black areas. As mentioned in the Introduction, the previously proposed FD formulation cannot account for such missing data regions, which complicates a quantitative SoS reconstruction. To avoid this drawback, the forward models in this study were implemented in space domain (SD). For this purpose, the measured spatial distribution of echo phase shift as well as the distribution of the to-be-reconstructed slowness are discretised on the same 2D Cartesian grid, as

$$\begin{aligned} &\Delta\sigma(\mathbf{r}') \rightarrow \Delta\sigma_{j,k} \\ &\Delta\Theta(\mathbf{r}', n) \rightarrow \Delta\Theta_{j,k}(n) \quad \text{(classical tracking)} \\ &\Delta\Theta(\mathbf{r}', n, m) \rightarrow \Delta\Theta_{j,k}(n, m) \quad \text{(CMA tracking)} \end{aligned} \tag{14}$$

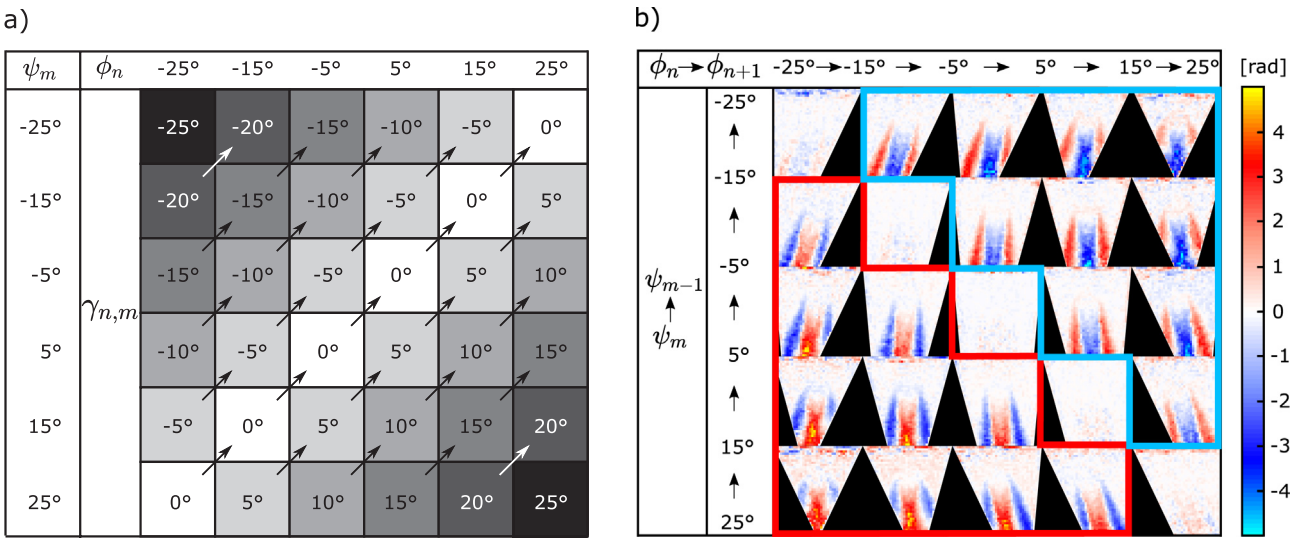

Fig. 4. (a) Combinations of Tx (ϕ_n), Rx (ψ_m) and common mid-angles ($\gamma_{n,m}$) used in this study. Tracking is performed between angle pairs having an identical mid-angle $\gamma_{n,m}$, indicated by arrows pointing along fields having the same gray value. No tracking was performed for $\gamma_{n,m} = \pm 25°$. (b) Exemplary phase shift maps from CMA tracking of the simple cylindrical inclusion phantom described in Materials and Methods.

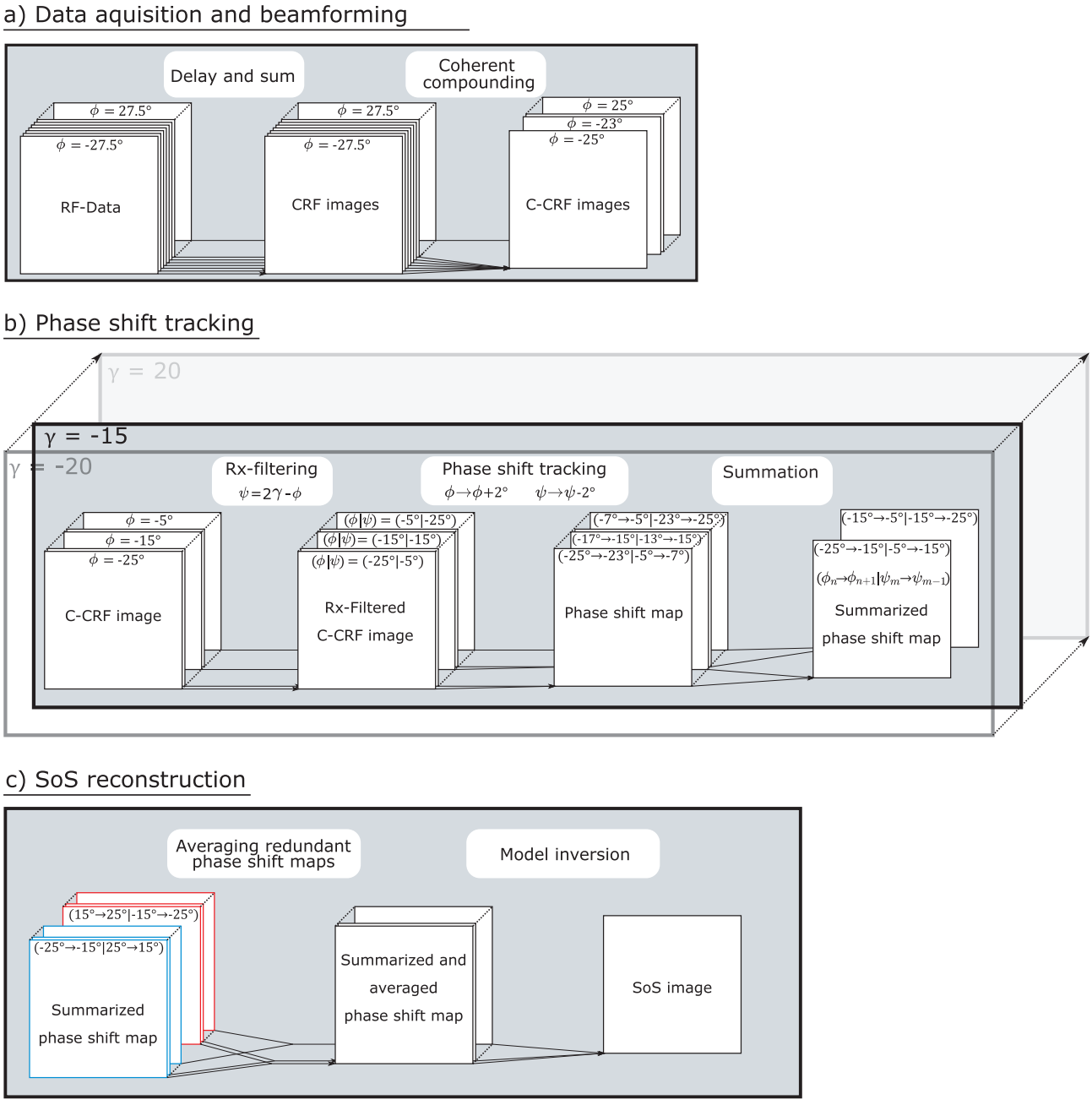

Fig. 5. Flow chart describing the working principle of CUTE, divided into three main parts.

Thereby, the number of nodes (j, k) in x- and z- direction is N_j and N_k, respectively. The nodes sample the chosen dimensions (X, Z) of the rectangular grid with a spatial resolution of $(\Delta x, \Delta z)$. Details are given further below.

The goal of the experimental part is to compare the old and the new model in a phantom study. Thus, both models are implemented as described by Eq. 15 and 17, respectively, derived from Eq. 6 and 13. In addition, we want to investigate whether taking into account the echo position offset is required for accurate phase shift predictions and SoS reconstruction, or whether CMA alone is already sufficient. For this purpose, we also implement an intermediate forward model where only CMA tracking - but not the echo position offset - is taken into account. This intermediate model is henceforth termed 'old CMA model' (Eq. 16).

- **Old model**

$$\Delta\Theta_{j,k}(n) = 2\pi f_0 \left[\sum_{j',k'} w^{n+1}_{j,k,j',k'} \Delta\sigma_{j',k'} - \sum_{j',k'} w^{n}_{j,k,j',k'} \Delta\sigma_{j',k'} \right] \tag{15}$$

- **Old CMA model**

$$\Delta\Theta_{j,k}(n, m) = 2\pi f_0 \left[\left(\sum_{j',k'} w^{n+1}_{j,k,j',k'} \Delta\sigma_{j',k'} + \sum_{j',k'} w^{m-1}_{j,k,j',k'} \Delta\sigma_{j',k'} \right) - \left(\sum_{j',k'} w^{n}_{j,k,j',k'} \sigma_{j',k'} + \sum_{j',k'} w^{m}_{j,k,j',k'} \sigma_{j',k'} \right) \right] \tag{16}$$

- **New model**

$$\Delta\Theta_{j,k}(n,m)= \frac{2\pi f_0}{\Omega'}\left(\sum_{j',k'} w^{n+1}_{j,k,j',k'}\Delta\sigma_{j',k'} + \sum_{j',k'} w^{m-1}_{j,k,j',k'}\Delta\sigma_{j',k'}\right) - \frac{2\pi f_0}{\Omega}\left(\sum_{j',k'} w^{n}_{j,k,j',k'}\Delta\sigma_{j',k'} + \sum_{j',k'} w^{m}_{j,k,j',k'}\Delta\sigma_{j',k'}\right) \quad (17)$$

with,

$$\Omega' = \cos\left[\tfrac{1}{2}(\phi_{n+1} - \psi_{m-1})\right] \qquad \Omega = \cos\left[\tfrac{1}{2}(\phi_n - \psi_m)\right] \quad (18)$$

Thereby the $w^{l}_{j,k,j',k'}$ are the integration weights defining the discrete line integrals along the US propagation paths defined by the angles Φ_l.

Any of the forward models in Eqs. (15)–(17) can be written in matrix notation, by vectorizing $\Delta\sigma$ and $\Delta\Theta$:

$$\Delta\theta = \boldsymbol{M}\Delta\boldsymbol{\sigma} \quad (19)$$

The system matrix $\boldsymbol{M}$ also takes into account the missing data regions (shadows) mentioned before, by rows containing zeros that project to zero values in the data vector $\Delta\theta$. Note that, even though all phase shift maps have areas of missing data, these areas are different for different angle steps so that each pixel of the chosen image area is covered by at least one of the phase shift maps. Therefore the SoS can be reconstructed in any pixel of the chosen image area. To reconstruct the values $\Delta\sigma'_{j',k'}$ from the measurements $\Delta\theta_{j,k}$, a Tikhonov pseudo-inverse of $\boldsymbol{M}$ is used:

$$\Delta\boldsymbol{\sigma}' = \boldsymbol{M}^{\mathrm{inv}}\Delta\theta \quad (20)$$

$$\boldsymbol{M}^{\mathrm{inv}} = (\boldsymbol{M}^T\boldsymbol{M} + \gamma_x\boldsymbol{D}_x^T\boldsymbol{D}_x + \gamma_z\boldsymbol{D}_z^T\boldsymbol{D}_z)^{\mathrm{inv}}\boldsymbol{M}^T \quad (21)$$

which minimizes the expression

$$C(\Delta\boldsymbol{\sigma}') = \|\Delta\theta - \boldsymbol{M}\Delta\boldsymbol{\sigma}'\|_2^2 + \|\gamma_x\boldsymbol{D}_x\Delta\boldsymbol{\sigma}' + \gamma_z\boldsymbol{D}_z\Delta\boldsymbol{\sigma}'\|_2^2 \quad (22)$$

Thereby, $\boldsymbol{D}_x$ and $\boldsymbol{D}_z$ are finite difference operators in x and z, respectively, and γ_x and γ_z are regularisation parameters. Regularisation of the spatial gradient of the slowness deviation enforces a smooth slowness profile without imposing a constraint on the mean SoS. The reconstructed SoS $c'_{j',k'}$ is finally recovered from the reconstructed slowness deviation $\Delta\sigma'_{j',k'}$ according to:

$$c'_{j',k'} = \left(\Delta\sigma'_{j',k'} + \frac{1}{\hat{c}}\right)^{-1} \quad (23)$$

3.2. Data acquisition and beamforming

Fig. 5 displays a flow chart that illustrates the various steps that finally lead to the SoS images. We used an L7-4 linear vascular probe (ATL Philips, WA, USA) for pulse-echo signal acquisition. This probe features 128 elements at a 0.29 mm pitch (resulting in 38.4 mm aperture length), and a bandwidth from 4 to 7 MHz with 5 MHz centre frequency. It was connected to a Vantage 64 LE (Verasonics Inc., WA, USA) for data acquisition. This research ultrasound system allows simultaneous ultrasound transmission on 128 elements and parallel digitisation of signals on 64 elements at a time, with real time data transfer via a PCI Express link to a host computer (Intel Core I7-4790 CPU and 16 GB RAM) for further processing. We implemented a dedicated scan sequence for the acquisition of plane wave pulse-echo data. As mentioned in the theory section, Tx focusing is required to achieve a lateral resolution of the phase shift maps. In this study, this was achieved via coherent plane-wave compounding [46,47] of plane wave acquisitions with closely spaced Tx angles, for which we chose a 0.5° angle resolution. To use the available angular aperture of the probe (grating lobes were inhibitive outside ± 30°), plane wave data was thus acquired with Tx angles ranging from −27.5° to 27.5° in 0.5° steps. To use all 128 elements on receive, echoes were recorded twice for each Tx angle, once on each of two 64-element sub-arrays, and then combined to a single data frame per angle. The 3D data set resulting from combining the 2D frames of all Tx angles is indicated in Fig. 5 (a).

Complex radio frequence (crf)-mode images (see definition in Theory section) were reconstructed off-line for each Tx angle from the raw crf data, using a delay-and-sum (DAS) algorithm:

$$u\left(x, z, \phi\right) = \sum_{el} \left[\alpha(x, z, el)\cdot s(\hat{t}_0(x, z, \phi, el))\right] \quad (24)$$

At each grid node with coordinates (x, z), the anticipated round-trip times $\hat{t}_0(x, z, \phi, el)$ to each receiving element el were calculated based on the anticipated SoS $\hat{c}$.

$$\hat{t}_0(x, z, \phi, el) = \frac{1}{\hat{c}}\left[\sin(\phi)(x - \tan(\phi)z) + \cos(\phi)z + \sqrt{(x_{el} - x)^2 + z^2}\right] \quad (25)$$

The apodization weights $\alpha(x, z, el)$ define the receive angular aperture within the limits of the aperture length. They were chosen so as to use a receive angular aperture of ±30° (limited by the grating lobes outside these angles), by setting them to 1 inside this aperture and to 0 outside. Crf-mode amplitudes were calculated on a rectangular grid, with dimensions 38.4 mm in x (aperture length) by 40 mm in z direction and with spatial resolution $dx = 0.29$ mm in x (pitch) by $dz = 0.037$ mm in z. The DAS step is indicated in Fig. 5 (a) ("delay and sum"), where each data frame is transferred to the corresponding crf-mode image.

As mentioned, the target Tx and Rx angle step size chosen for SoS reconstruction was 10°. Given the SoS contrast of the phantoms used in this study, however, the magnitude of the aberration delay change $\Delta\theta$ obtained with this step size resulted in phase shifts above π and thus in phase aliasing when quantifying the phase shift according to Eq. 5 and 8. By inspecting the phase shift maps, we empirically determined that a tracking angle step size of 2° was just sufficiently small to avoid phase aliasing in all phantoms. Therefore, from the previously reconstructed crf-mode images, we generated images for tracking Tx angles $\phi \in \{-25°: 2°: 25°\}$. For synthetically focused Tx steering along any of these tracking angles ϕ, the 0.5°-spaced crf-mode images were coherently compounded using a Gaussian weighting function centered at the respective tracking angle with a Gaussian radius of 2.5°. This step is indicated in Fig. 5 (a) ("coherent compounding"), where a multitude of crf-images (e.g. for ϕ ranging from −27.5° to −22.5°) is transferred to a single coherently compounded crf ("c-crf") image (e.g. at −25°).

For the old (non-CMA) model, these c-crf images were directly used for tracking. For the CMA tracking, beamformed images are required for different Rx steering angles per each Tx steering angle. This could in principle be implemented via the apodization weights in Eq. 24, but we instead used the following approach, which was implemented directly as part of the tracking algorithm: for one tracking Tx angle $\phi \in \{-25°: 2°: 25°\}$ at a time, the full-Rx-aperture c-crf mode image was transformed to a sequence of synthetically Rx-steered images via spatial filtering, for Rx steering angles corresponding to the different mid-angles $\gamma_{n,m} \in \{-20°: 5°: 20°\}$. For the spatial filtering, the respective c-crf-mode image was transferred to k-space by calculating the 2D discrete Fourier transform. In k-space, a specific pair $(\phi|\psi)$ of Tx and Rx angles detects a line of the tissue's reflectivity function along the mid-angle $0.5(\phi + \psi)$. Conversely, given the Tx angle ϕ and a target mid-angle $\gamma_{n,m}$, filtering the k-space for a line along $\gamma_{n,m}$ automatically results in the k-space representation of a c-crf-image as if acquired with a tracking Rx angle $\psi = 2\gamma_{n,m} - \phi$.

Filtering for a sharply defined $\gamma_{n,m}$ would result in an image as if acquired with zero angular aperture that would provide no spatial resolution perpendicular to $\gamma_{n,m}$. Rather than filtering for only one angle, an angle-dependent Gaussian weighting function was thus used with a radius of ± 1.25° to synthesise an Rx angular aperture radius of 2.5°. The spatial frequency spectrum was multiplied with a set of such weighting functions centered at the different mid-angles $\gamma_{n,m}$, and then inverse

Fourier transformed, to obtain a set of synthetically Rx-filtered c-crf images corresponding to the different $\gamma_{n,m}$. In comparison to a space-domain Rx apodization approach, the k-space approach has the advantage that the Rx beamforming is more similar to the Tx beamforming: k-space filtering is conceptually similar to using "receiving plane waves" for echo detection, and synthesising an Rx angular aperture in k-space is similar to coherent plane-wave compounding in receive, thus improving the Tx/Rx reciprocity in comparison to an element-based Rx beamforming. The spatial frequency filtering step is indicated in Fig. 5 (b) ("Rx-filtering"), exemplarily for mid-angle -15°, where an Rx-filtered c-crf image is generated from each full-Rx-aperture c-crf-image that is used for tracking around this mid-angle.

The bandpass to enforce a constant Λ according to Eq. 12 was also implemented as part of the spatial frequency filtering: for each $\gamma_{n,m}$, the already Rx-filtered spectrum was multiplied with a k-dependent Gaussian weighting function (where k is the modulus of the spatial frequency vectors) with empirically determined (ψ, $\gamma_{n,m}$)-dependent center point and radius, to enforce a constant center k of the final filtered spectrum.

As mentioned earlier, it is the beamforming that provides the spatial resolution to CUTE. More precisely, the lower resolution limit is determined by the resolution of the crf images (whereas the final resolution is determined also by phase tracking kernel size and SoS regularization). We have performed simulations (not shown) to determine the spatial resolution of the c-crf (Tx-only tracking) and the Rx-filtered c-crf images (CMA tracking), given the chosen beamforming parameters: in Tx-only tracking where the full available Rx angular aperture is used, the resolution limit is determined by the Rx angular aperture to 0.3 mm (FWHM). In CMA tracking, the resolution limit results from the combination of the 2.5° Tx and Rx angular aperture radii, to 2.3 mm (FWHM).

3.3. Phase shift tracking

Based on the c-crf images, the echo phase shift was finally determined according to Eq. 5 (Tx-only tracking) and Eq. 8 (CMA tracking). For Tx-only tracking, phase shift maps were calculated between the full-Rx-aperture c-crf images of successive Tx angles that were spaced by the 2° Tx angle spacing. CMA tracking was performed as follows: looping through successive Tx angles, the Rx filtering was applied to the full-Rx-aperture c-crf image of a respective actual Tx angle at a time. For each $\gamma_{n,m}$, a phase shift map was calculated between the respective previous Rx-filtered c-crf image and the respective actual Rx-filtered c-crf image corresponding to that $\gamma_{n,m}$. The resulting phase shift maps were stored for accumulation (see further below), and the actual Rx-filtered c-crf images were stored to serve as the previous Rx-filtered c-crf images in the next tracking step. For both Tx-only and CMA tracking, the tracking kernel size was chosen $\Delta x = 2$ mm and $\Delta z = 2$ mm. For CMA tracking, this processing step is indicated in Fig. 5 (b) ("phase shift tracking"), where two successive Rx-filtered c-crf images (e.g. from angle pair $(-25°|-23°)$ to angle pair $(-5°|-7°)$) are transferred to one phase shift map. Phase shift maps were then summed over successive 2° angle steps to obtain the echo phase shift over 10° steps between Tx angles ϕ_n and ϕ_{n+1} (Tx-only tracking) and Tx/Rx angle pairs $(\phi_n|\psi_m)$ and $(\phi_{n+1}|\psi_{m-1})$ (CMA tracking). For CMA tracking, this is exemplarily shown in Fig. 5 (b) ("summation") for mid-angle -15°, where phase shift maps from angle pair (-25° | −5°) to angle pair (-15° | −15°) are summed to a single final phase shift map.

3.4. SoS reconstruction

For SoS reconstruction, redundant phase shift maps were averaged ("averaging redundant phase shift maps" in Fig. 5 (c)), and the result downsampled onto a Cartesian grid with $N_x = 40$ by $N_z = 40$ grid nodes covering $X = 38.4$ mm by $Z = 40$ mm with a resolution of $\Delta x = 0.96$ mm by $\Delta z = 1$ mm. The forward models according to Eq. 19 and 20 were correspondingly formulated for the same grid resolution, resulting in a system matrix with (40·40) by (40·40) elements per 10° angle step, totalling 8000 by 1600 elements for Tx-only tracking (5 Tx angle steps) and 16000 by 1600 elements for CMA tracking (10 combined Tx/Rx angle steps). For line integration, bi-linear interpolation weights were chosen.

The regularization parameters γ_x and γ_z are subject to a trade-off between reducing artefacts (by enforcing a smooth slowness profile) and maximising spatial resolution. For this study, they were chosen so as to clearly distinguish the different phantom compartments (Old model: $\gamma_x = 2.33$, $\gamma_z = 0.23$; old CMA model: $\gamma_x = 5.68$, $\gamma_z = 0.11$; New model: $\gamma_x = 9.09$, $\gamma_z = 0.18$).

The pseudo-inverses were pre-calculated according to Eq. 21 and stored. The pre-calculated pseudo-inverses were then applied to the vectorized phase shift data for reconstructing the slowness distribution, and – in turn – the SoS. This step (application of pesudo-inverse, calculation of SoS from slowness) is indicated in Fig. 5 (c), "model inversion". With the chosen grid resolution, the matrix multiplication takes a fraction of a second (e.g. 0.04 s on the used Intel Core I7-4790 CPU and 16 GB RAM) when implemented in Matlab®.

3.5. Phantom design and materials

For this study, two different sets of phantoms were designed and investigated. The first set of phantoms was designed with the goal to represent particular geometries of the anatomical structure of the abdominal wall and liver, since liver imaging is one of our envisaged clinical applications. To mimic focal lesions (as e.g. for the diagnosis of breast cancer) the phantoms of the second set contain cylindrical inclusions. The geometries of the spatial distribution of SoS in the different phantoms are shown in Fig. 6. Note that, in the following, even though the absolute SoS values of the different phantom compartments may deviate from real tissue depending on literature references, we have taken care to chose realistic SoS contrasts between the different compartments.

- Liver mimicking phantoms
  - (a) *Two-layer phantom*: This phantom contained two horizontal (parallel to x) layers, mimicking a single fat layer (F1: 1420 ms^{-1}) on top of liver tissue (L: 1555 ms^{-1}). The special feature of this type is the complete absence of lateral SoS variations but a pronounced axial variation. This geometry occurs when imaging the liver sagittally through the linea alba.
  - (b) *Four-layer phantom*: In comparison to the two-layer phantom,

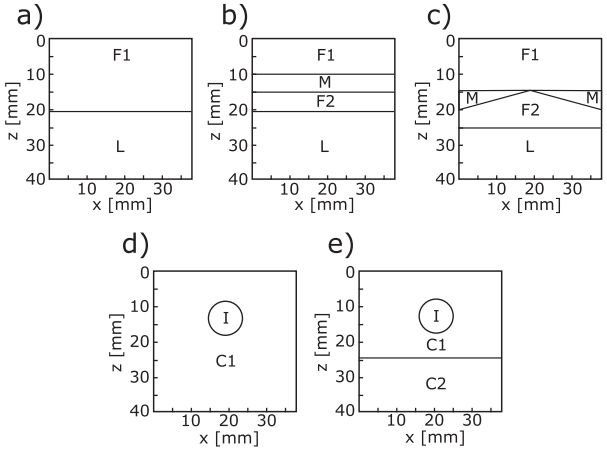

Fig. 6. Sketch of the phantom geometries. The reference SoS were determined using a through transmission time-of-flight set-up with an accuracy of ±5 ms^{-1}, to F1: 1420 ms^{-1}, F2: 1420 ms^{-1}, M: 1555 ms^{-1}, L: 1555 ms^{-1}, C1: 1540 ms^{-1}, C2: 1570 ms^{-1}, I: 1570 ms^{-1}.

two additional layers were added, mimicking the rectus abdominis muscle (M: 1555 ms^{-1}) and the post-peritoneal fat layer (F2: 1555 ms^{-1}), representative e.g. for a sagittal section lateral to the linea alba. The purpose of this phantom was to compare the different models in a layer structure with axial variations near the axial resolution limit.

(c) *Laterally varying muscle diameter (LVMD) phantom*: In contrast to the four-layer phantom, the rectus abdominis (M: 1555 ms^{-1}) layer deviates in this phantom from the parallel layer structure and consists of two wedge-shaped areas, as when imaging the liver in a transverse section through the linea alba. This phantom thus provides a moderate lateral variation of SoS in addition to the axial variation.

- Cylindrical inclusion phantoms

(d) *Simple inclusion phantom*: This phantom was composed of a uniform background compartment (C1: 1540 ms^{-1}) containing a cylindrical inclusion (I: 1570 ms^{-1}).

(e) *Two-layer phantom with inclusion*: In comparison to the simple inclusion phantom, the background compartment of this phantom contained an additional layer (C2: 1570 ms^{-1}) below the inclusion.

The phantom components were produced based on porcine gelatine (Geistlich Spezial Gelatine, health and life AG, Switzerland). In a first step, gelatine was dissolved in 75 °C H_2O with various weight contents for the different phantom compartments (C1: 15 wt%, M, L: 20 wt%, C2, I: 25 wt%). To provide a uniform diffuse echogenicity, various weight contents of cellulose (Sigmacell Cellulose Type 20, Sigma Aldrich, Switzerland) were added (C1: 1 wt%, M,L, C2: 2 wt%, I: 0.5 wt %). The resulting SoS of the different gelatin solutions after gelling were: C1: 1540 ms^{-1}, M, L: 1555 ms^{-1}, C2, I: 1570 ms^{-1}. For the fat mimicking compartments (F1, F2) in the liver mimicking phantoms, oil-in-gelatin emulsions were produced [48–51]. For this purpose, medium-chain triglycerides oil (Ceres-MCT Oil, Puravita, Switzerland) (SoS = 1350 m/s) was slowly blended under continuous stirring into the still hot aqueous gelatine base solution (20 wt% gelatine and 2 wt% cellulose) using a Visco Jet cone-stirrer (VISCO JET Agitation Systems, Germany). During this process, small oil droplets were formed and captured via hydrophobic interaction by the lipophilic part of the gelatine strings. After cooling the emulsion, the oil droplets were trapped within the gelatine matrix. The resulting SoS was determined by the emulsion's relative MCT oil weight content, in this study 0.65 wt%, resulting in a SoS of 1420 m/s. The mentioned SoS values were determined using a through-transmission time-of-flight set-up (accuracy ±5 m/s) and will serve in the following as reference for comparing the accuracy of phase shift predictions and reconstructed SoS values.

4. Results

4.1. Echo phase shift

Similar to what has already been done for the old model (see. Fig. 1), Fig. 7 shows profiles of the measured phase shifts and of the predictions made by the old CMA (Eq. 16) and the new model (Eq. 17) in the same two phantoms, for the Tx-Rx tracking combination of $\phi_n \rightarrow \phi_{n+1} = 5^\circ \rightarrow 15^\circ$ and $\psi_m \rightarrow \psi_{m-1} = -5^\circ \rightarrow -15^\circ$. In case of the four-layer phantom, the axial profile at $x = 19.2$ mm is shown. For the simple cylindrical inclusion phantom, the lateral profile at $z = 30$ mm is shown. For a more quantitative comparison, Table 1 summarizes the root-mean-square error (RMSE) of the complete phase shift datasets for all models according to:

$$\mathrm{RMSE}_{\Delta\Theta} = \left(\frac{\sum_{n=1}^{5}\sum_{m=1}^{5}\sum_{j=1,k=1}^{N_j,N_k} [\Delta\Theta^{\star}_{j,k}(n,m) - \Delta\Theta_{j,k}(n,m)]^2}{N_j \cdot N_k \cdot 5 \cdot 5} \right)^{-1/2} \quad (26)$$

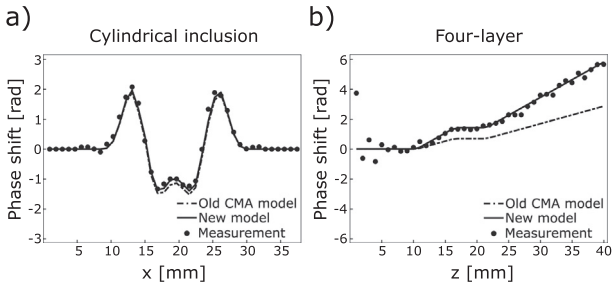

Fig. 7. Comparison of the measured phase shift to the predictions made by the old CMA model and by the new model, for Tx-Rx angle steps of $\phi_n \rightarrow \phi_{n+1} = 5^\circ \rightarrow 15^\circ$ and $\psi_m \rightarrow \psi_{m-1} = -5^\circ \rightarrow -15^\circ$, around the mid-angle $\gamma_{n,m} = 0^\circ$. (a) Lateral profile ($z = 30$ mm) of the simple cylindrical inclusion phantom. (b) Axial profile ($x = 19.2$ mm) of the four-layer phantom.

Table 1
$\mathrm{RMSE}_{\Delta\Theta}$ of the measure phase shift in comparison to the different forward models in two exemplary phantoms.

	$\mathrm{RMSE}_{\Delta\Theta}$ [rad]	
	Four-layer	Cylindrical inclusion
Old Model	0.738	0.423
CMA Model	0.852	0.339
New Model	0.384	0.320

where $\Delta\Theta^{\star}$ denotes the measured phase shift and $\Delta\Theta$ the phase shift predicted by the forward model at grid node (j, k) for all combinations of Tx/Rx pairs $(\phi_n|\psi_m)$ that are outlined in Fig. 4(b).

In case of the simple cylindrical inclusion phantom (Fig. 7 a), the phase shift that is predicted by the old CMA model matches the measured data already well. Note especially that it not only predicts the correct phase shift amplitude but also the spatial profile, whereas the latter was wrongly predicted with the old model (compare to Fig. 1). The improved accuracy is confirmed by the $\mathrm{RMSE}_{\Delta\Theta}$. These observations indicate that the wrong prediction of the profile by the old model was due to an implicit Rx-angle dependence of the real data that was not accounted for. By explicitly modelling this dependence in the CMA approach, a much more accurate prediction was obtained.

In comparison to the old CMA model, the new model gives only a slight improvement for the cylindrical inclusion (seen both in Fig. 7 and in the $\mathrm{RMSE}_{\Delta\Theta}$). This changes, however, completely for the four-layer phantom (Fig. 7 b): here the old CMA model underestimates the measured phase shift by a factor two similar to what was already observed for the old model in Fig. 1, demonstrating that the CMA approach alone cannot describe the reality well. When, however, accounting also for the echo position offset (new model), the predicted phase shift matches the measured data very well. The $\mathrm{RMSE}_{\Delta\Theta}$ again confirms that the new model describes the reality best.

Note that, in both phantoms, the differences in the $\mathrm{RMSE}_{\Delta\Theta}$ between the models appear less extreme than one could expect based on the comparison of the phase shift profiles: the $\mathrm{RMSE}_{\Delta\Theta}$ is reduced by roughly a factor two for the four-layer phantom and 1.3 for the cylindrical inclusion phantom between the old and the new model, whereas the profiles show perfect agreement for the new model. The reason is following: profiles are shown for angle combinations where the phase shift magnitude is large to provide a good SNR. For many other angle combinations, the phase shift magnitude is smaller, and over all the $\mathrm{RMSE}_{\Delta\Theta}$ is to a large part determined by phase noise. It therefore also comes at no surprise that the $\mathrm{RMSE}_{\Delta\Theta}$ of the new model are very similar for both phantoms, as they represent the phase noise level rather than an actual deviation between prediction and experiment.

Fig. 7 together with the $\mathrm{RMSE}_{\Delta\Theta}$ thus demonstrate that both

features, the CMA approach and the modelling of the echo position error, are a minimum requirement for an accurate forward model and thus for achieving quantitative SoS imaging in CUTE.

4.2. SoS images

The comparison between the measured phase shifts and the predictions made by the forward models on its own already clearly falsifies the general applicability of the old and of the old CMA model, since both fail in predicting the correct phase shift in the four-layer phantom.

In a next step, we want to investigate whether the two proposed changes to the CUTE forward model together are not only a minimum requirement, but also sufficient for achieving quantitative SoS results. This investigation has to be performed based on SoS images as opposed to phase shift predictions for following reason: subtle differences in phase shift predictions of a forward model can unpredictably affect the reconstructed SoS, given the ill-posed nature of the inverse problem. In case of focal lesions, for example, one might argue that the CMA approach alone is sufficiently accurate, but we will see that this is not true for the SoS images. Similarly, small errors in the proposed new model could lead to unexpected large errors in the reconstructed SoS that obliterate the usefullness of the new model.

Therefore, to further compare the different models and evaluate their performance, the reconstructed SoS images are shown for all five phantoms in Figs. 8–12. The color scale was chosen in a trade-off between covering a large SoS range and allowing to perceive small SoS variations (e.g. artifacts). As a consequence, image regions in which the SoS deviates by $\gtrsim$70 ms^{-1} from the true upper and lower SoS of the liver mimicking phantom appear saturated. To allow a fair comparison between the images of the two phantom sets (e.g. the level of artifacts), the same color scale was also used for the cylindrical inclusion phantoms.

Note that SoS reconstruction results are shown for various different choices of the *a priori* SoS $\hat{c}$. As described in Materials and Methods, $\hat{c}$ is used for the reconstruction of the crf-mode images. In an ideal scenario, the deviation of the actual SoS from this value results in an echo phase shift and a reconstructed slowness deviation that – together with $\hat{c}$ – finally leads to the correct SoS independent of $\hat{c}$ according to Eq. 23. In reality, however, errors can be introduced via the influence of $\hat{c}$ on the beamforming (i.e. distortions of the spatial pulse shape as well as propagation angle) and/or via errors of the SoS reconstruction model. While being a potential source of errors, $\hat{c}$ is also a parameter that would be conveniently accessible to manipulation by a radiologist. For this reason, we examine the sensitivity of the SoS results towards variations of $\hat{c}$, by comparing results for different $\hat{c}$ covering a range of 1420 ms^{-1} to 1555 ms^{-1} for the liver mimicking phantoms and 1480 ms^{-1} to 1570 ms^{-1} for the inclusion phantoms.

To aid a quantitative evaluation and the comparison of the different models, we provide several metrics of the SoS values: the rightmost columns of Figs. 8–12 show the mean and standard deviation of the reconstructed SoS within each individual compartment over the full respective compartment area. In addition, Fig. 13 shows the root-mean-square errors (RMSE) of all phantoms according to:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{j,k} (c_{j,k}{}' - c_{j,k})^2}{N_j \cdot N_k}} \tag{27}$$

where $c'_{j,k}$ is the reconstructed SoS and $c_{j,k}$ the ground-truth SoS at the grid nodes (j, k). Note that, independent of the phantom geometry, the *a priori* SoS $\hat{c}$ or the forward model, all SoS images show artifacts in the top few millimetres (see Figs. 8–12). These artifacts originate from erroneous values (outliers) in the top few millimetres of the phase shift maps (see e.g. Fig. 1, 4 and 7). We hypothesize that these errors result from element cross-talk during the plane wave transmission: whenever an element is activated, all other elements instantaneously detect an electric interference. The larger the Tx angle, the more the interferences from all transmitting elements are stretched over time and obscure superficial echoes. As cross-talk related artifacts are not related to the forward model, the top 5 mm of the SoS images were excluded from the calculation of the SoS metrics.

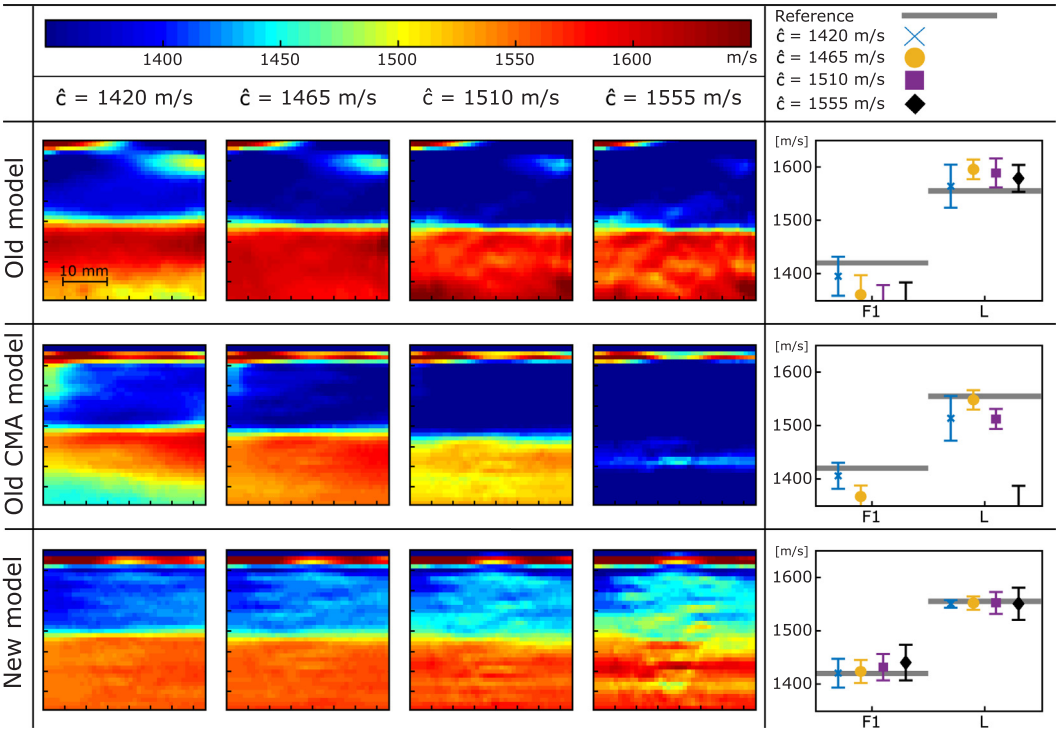

Fig. 8. SoS images of the two-layer phantom mimicking the liver (SoS = 1555 m/s) and a single fat layer (SoS = 1420 m/s) (from posterior to anterior). The SoS images were reconstructed based on the different forward models (rows) and for different *a priori* SoS values $\hat{c}$ (columns). The SoS images are sampled on 40 (x) by 40 (z) grid nodes covering 38.4 mm (x) by 40 mm (z).

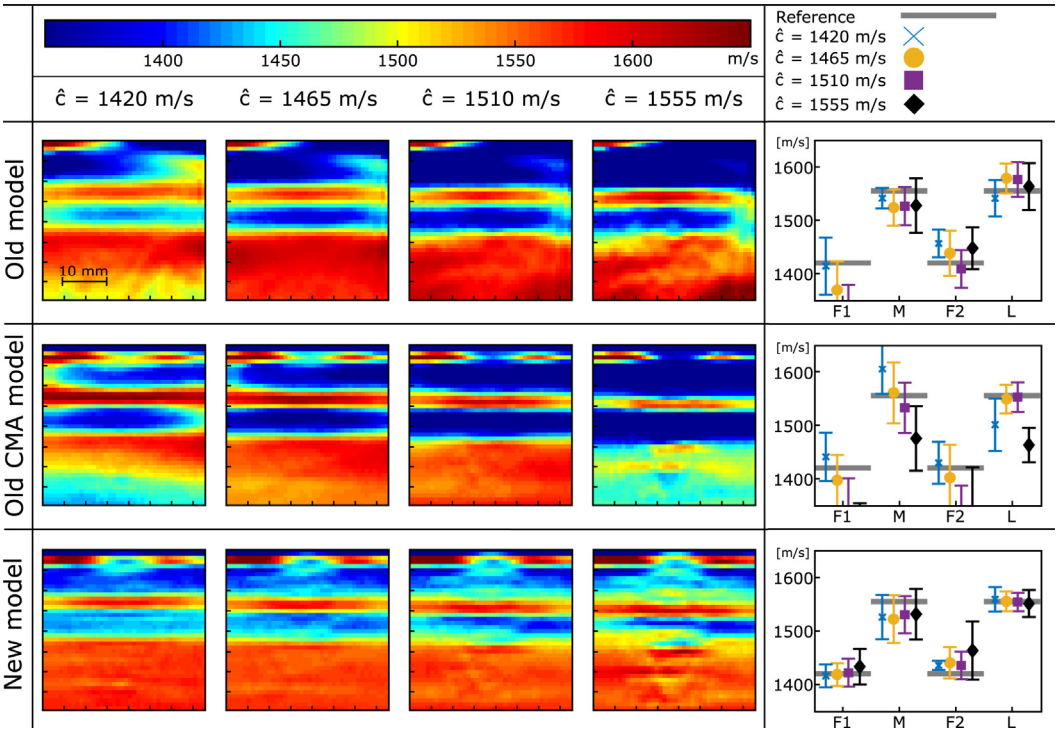

Fig. 9. SoS images of the four-layer phantom mimicking the liver (SoS = 1555 m/s), a postperitoneal layer (SoS = 1420 m/s), the rectus abdomins (SoS = 1555 m/s) and the subcutaneous fat layer (SoS = 1420 m/s) (from posterior to anterior). The SoS images were reconstructed based on the different forward models (rows) and for different *a priori* SoS values $\hat{c}$ (columns). The SoS images are sampled on 40 (x) by 40 (z) grid nodes covering 38.4 mm (x) by 40 mm (z).

4.2.1. Old model

For intermediate $\hat{c}$, the old model leads to SoS images of the two- and four-layer phantoms (see Figs. 8 and 9 - top rows) that seem reasonable in terms of a uniform spatial distribution of SoS in the liver-mimicking compartment L. At the lower limit of the chosen range of $\hat{c}$, however, the SoS inside L becomes less uniform, and the contrast between the compartments varies with $\hat{c}$. In spite of these shortcomings, one could argue, the images of these two phantoms look quite

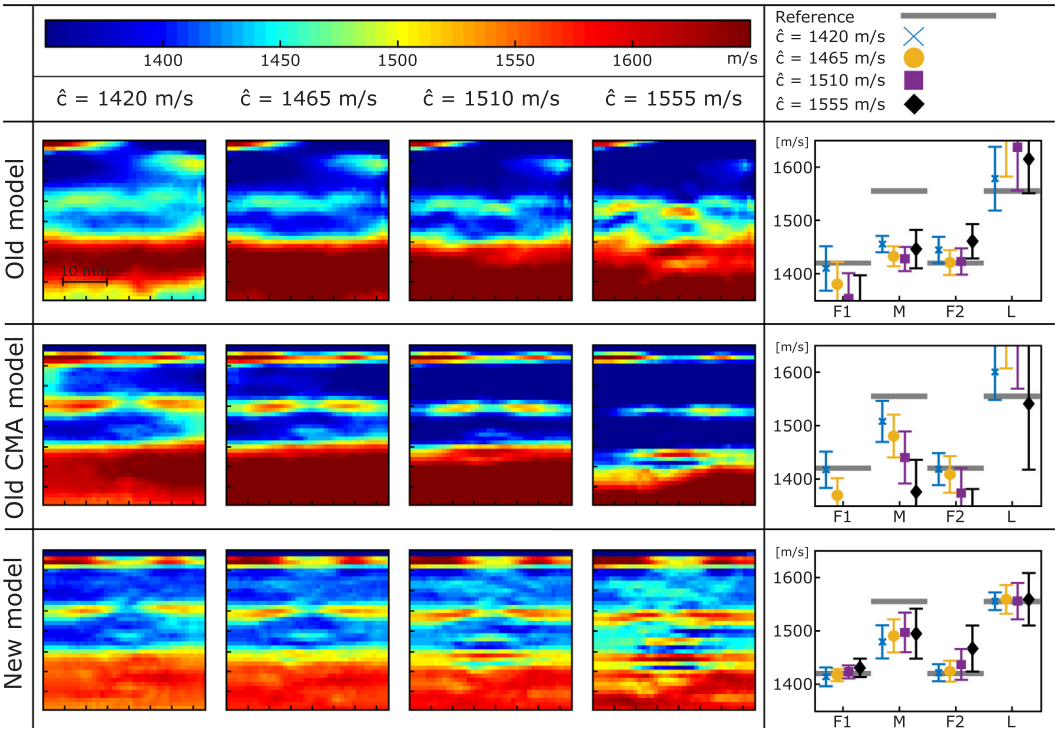

Fig. 10. SoS images of the LVMD phantom mimicking the liver (SoS = 1555 m/s) covered by a triangular postperitoneal fat layer (SoS = 1420 m/s), the rectus abdomins (SoS = 1555 m/s) and the subcutaneous fat layer (SoS = 1420 m/s) (from posterior to anterior). The SoS images were reconstructed based on the different forward models (rows) and for different *a priori* SoS values $\hat{c}$ (columns). The SoS images are sampled on 40 (x) by 40 (z) grid nodes covering 38.4 mm (x) by 40 mm (z).

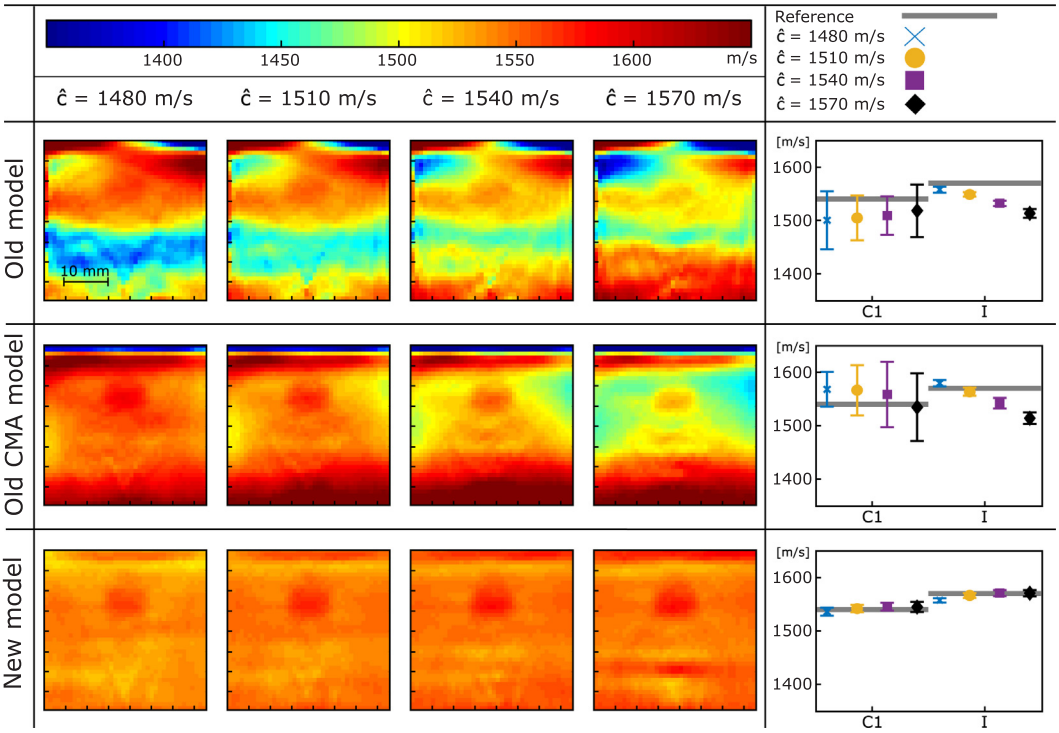

Fig. 11. SoS images of the simple cylindrical inclusion phantom. The SoS images were reconstructed based on the different forward models (rows) and for different *a priori* SoS values $\hat{c}$ (columns). The SoS images are sampled on 40 (x) by 40 (z) grid nodes covering 38.4 mm (x) by 40 mm (z).

reasonable. This changes, however, dramatically for the LVDM phantom (Fig. 10 - top row): here the SoS in compartments L is substantially overestimated, and heavily non-uniform. The SoS in compartments M is underestimated, and the spatial distribution wrongly reconstructed, i.e. an artifactual high-SoS area appears between the wedge-shaped parts of M. Note that, in all phantoms, low spatial frequency variations of SoS appear in at least part of the image area. These variations are most clearly seen in the SoS images of the cylindrical inclusion phantoms (Figs. 11 and 12 - top rows), where they make it difficult – if not impossible – to distinguish the cylindrical inclusions from the background, and impossible to perceive the absence or presence of compartment C2.

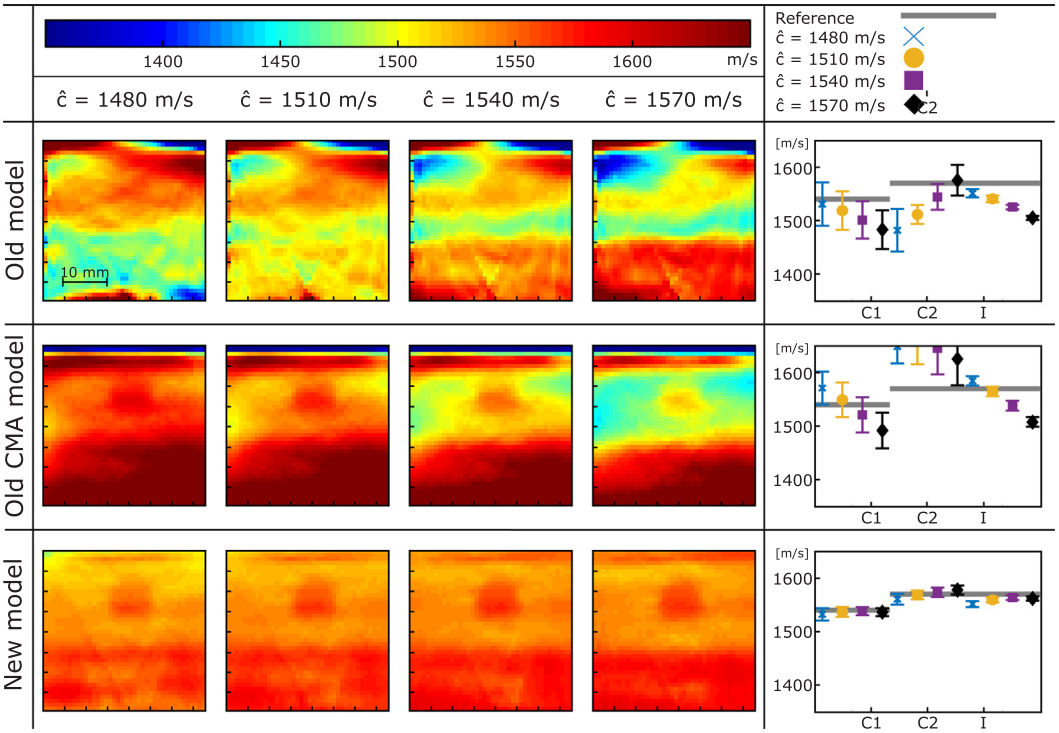

Fig. 12. SoS images of the two-layer cylindrical inclusion phantom. The SoS images were reconstructed based on the different forward models (rows) and for different *a priori* SoS values $\hat{c}$ (columns). The SoS images are sampled on 40 (x) by 40 (z) grid nodes covering 38.4 mm (x) by 40 mm (z).

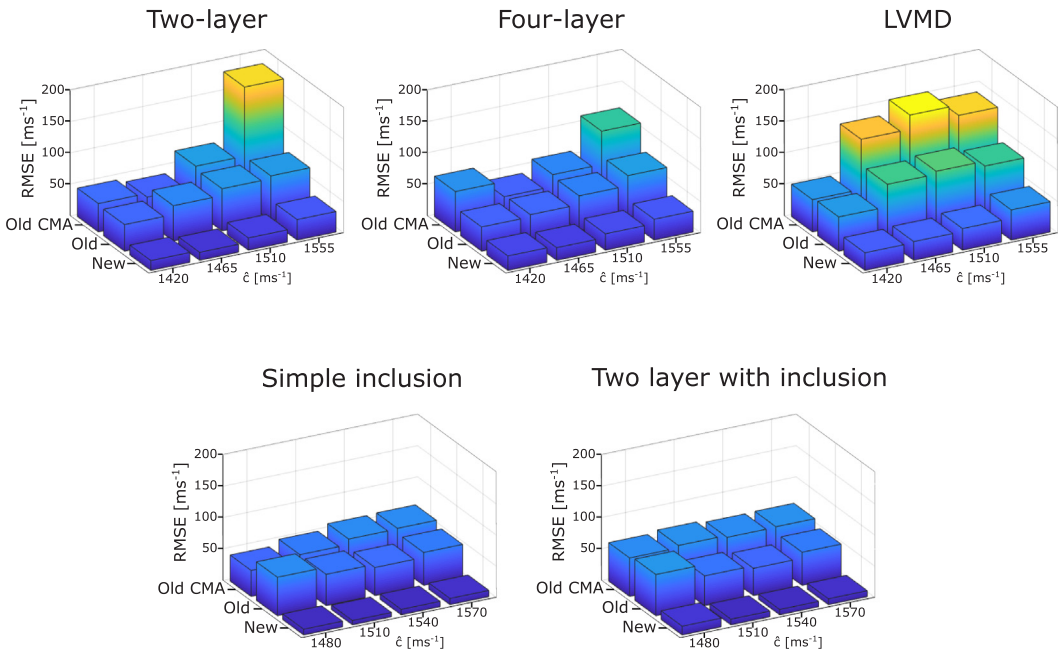

Fig. 13. Root-mean-square errors (RMSE) of all five phantoms, calculcated according to Eq. 27. For each phantom, the RMSE of the old CMA, the old and the new model at the various *a priori* SoS $\hat{c}$ are shown.

4.2.2. Old CMA model

In all phantoms (see Figs. 8–12 - middle rows), the old CMA model leads to even stronger low spatial frequency variations and a stronger dependence of the different compartments' SoS on $\hat{c}$. On the upside, it is better at reconstructing the true spatial profile of the SoS of the compartments M in the LVMD phantom. In the cylindrical inclusion phantoms (Figs. 11 and 12 - middle rows), the low spatial frequency variations are more symmetric than with the old model, and less dependent on $\hat{c}$, allowing a better perceptibility of the inclusions. Apart from that, the presence/absence of compartment C2 can still not be perceived well: due to the gradient of SoS in supposedly uniform areas, it is difficult to judge whether the apparent boundary at the depth of the surface of compartment C2 is an actual boundary or simply an artifact of the color scale.

4.2.3. New model

In comparison to the old and the old CMA model, the new model leads to strongly improved SoS images in the liver phantoms (Figs. 8–10 - bottom rows). Apart from some high-frequency artifacts for certain values of $\hat{c}$ (on which we comment further below) they show a much more uniform distribution of SoS in the different compartments, and this largely independent of $\hat{c}$. As already with the old CMA model, the compartments M in the LVMD phantom are well delineated. The average SoS values agree well with the references, this again independent of $\hat{c}$. The only exception is seen in the M and F2 compartments of the four-layer and the LVMD phantom, where the SoS is biased most probably due to the partial volume effect. In all liver phantoms, the uniformity of the L compartment and the resolution of its surface is best for $\hat{c}$ near the SoS of the overlaying F layers. With increasing $\hat{c}$, high spatial frequency artifacts appear and gradually increase at and inside the L compartment, and the resolution of the surface becomes worse. We suggest that these findings are caused by a decreasing beamforming quality at the depth of the L compartment – resulting in phase shift noise – due to the increasing deviation of the $\hat{c}$ from the actual SoS of the top compartments. Note that the high spatial frequency artifacts lead to fluctuations of SoS around a reasonable constant value rather than skewing the spatial profile of SoS within uniform layers.

Also in the cylindrical inclusion phantoms (Figs. 11 and 12 - bottom rows), the new model leads to SoS images that show the expected uniform distribution inside the different as well as the expected quantitative SoS values nearly independent of $\hat{c}$. The cylindrical inclusion as well as the absence or presence of the C2 compartment can now be clearly identified. High spatial frequency artifacts are seen to a lesser extent than in the liver phantoms. This can be explained by the fact that – on one hand – the chosen range of $\hat{c}$ is smaller than for the liver phantoms, and – on the other hand – the SoS of the first (C1) compartment is far from both extremes of the $\hat{c}$ range, so that beamforming is less affected for any of the $\hat{c}$ values. Note that, even though the old CMA model and the new model led to nearly identical phase shift profiles in Fig. 7, the resulting SoS images are very different. This demonstrates how – due to the ill-posedness of the inverse problem – small differences in phase shift predictions can cause unforeseen large differences in the final SoS images.

The superiority of the new model over the old and the old CMA model is further confirmed by the RMSE that are shown in Fig. 13: whereas the RMSE strongly depends on $\hat{c}$ with the old and the old CMA model, it is much more stable with the new model, and in all phantoms and for all $\hat{c}$, the new model clearly outperforms the old as well as the old CMA model.

5. Discussion and conclusion

Departing from the observation of wrong predictions of experimental echo phase shift by the old model, we proposed in this study two fundamental changes to the CUTE methodology: the CMA approach with an explicit formulation of the dependence of echo phase shift on both Tx and Rx angle, and the explicit consideration of the echo position offset in the forward model. The comparison of the measured and predicted phase shifts clearly demonstrates that only the new model leads to remarkable accuracy in all investigated phantom geometries. Both novel features together are thus a necessary prerequisite for an accurate forward model.

Due to the ill-posed nature of the inverse problem, however, it is difficult to predict how the differences in the forward models affect the final SoS images. Consequently, SoS images are needed to determine whether the new model is not only a necessary prerequisite for quantitative imaging but also – within the approximations made – a

sufficient one. The SoS images have shown that the new model outperforms the old and old CMA model in all phantoms in view of RMSE, level of artifacts as well as the stability against variations of the *a priori* assumed SoS $\hat{c}$. A slight positive dependence of reconstructed SoS on $\hat{c}$ is observed in all phantoms, which may be explained by an imperfect calibration of the beamforming or other factors not accounted for, such as the curvature in elevation of the probe's acoustic lens, but – given the small magnitude of the $\hat{c}$-influence compared to the standard deviations – this was considered beyond the scope of this study.

Given the ill-posed nature of the inverse problem, one may doubt whether the hyperparameters involved in the model implementations can have a similarly unpredictable influence as differences in phase shift predictions, and whether one may thus want to take conclusions based on SoS images with care. To avoid – as far as possible – the influence of hyperparameters on the results, we have explicitly restrained from optimizing hyperparameters for the different models separately but have chosen identical parameters wherever possible. More specifically, we have chosen: identical Tx angles and Tx angular apertures in all models, Rx angles and angular apertures identical to Tx angles and angular apertures in the CMA-based models, and an Rx angular aperture in the old model that is (nearly) identical to the Rx angle range in the CMA-based models. Only the regularisation parameters were chosen differently. They were, however, not optimized for the different models independently, but with the goal to provide SoS images of the chosen phantoms that are – to the eye – similar in terms of axial resolution for a fair comparison of artifact level. For this reason, we argue, the SoS images provide a fair comparison of model performance, and we conclude that the new model is not only the only one capable of correctly predicting echo phase shift among the investigated phantom geometries, but it is even able to correctly predict the small nuances that are needed to obtain a good SoS image, and this even given the limitations of the quite coarse approximations made (i.e. straight-ray propagation, negligible wave front distortion …).

A key assumption made in this study was the straight-ray approximation of US propagation, neglecting diffraction and refraction. Given the lateral resolution of the final SoS image (few mm) in relation to the wavelength (0.2 mm at 5 MHz), diffraction will have played a minor role. With the realistically high SoS contrast between the liver-mimicking phantoms' layers (135 m/s), one would expect that refraction had an important influence. In horizontally layered media, the effect of refraction is mainly a change of propagation angle which can bias the phase shift measurements. Based on the fact that correct quantitative results were obtained using the new model, one can conclude that such biases were negligible even for the LVMD phantom with the continuously varying diameter of the wedge-shaped compartments M. More influence of refraction would be expected for the cylindrical inclusion phantoms, where the inclusion poses a abrupt lateral variation of SoS. Our results suggest that also here, the straight-ray approximation was reasonable in the sense that correct quantitative results could be obtained, but given a substantially smaller SoS contrast (30 m/s) than in the LVMD phantom (135 m/s). Independent experiments have shown that – for the same size of the inclusion – an increasing SoS contrast leads to an increasing level of artifacts below the inclusion, as well as to an underestimation of the SoS inside the inclusion. Such artifacts can on one hand be explained by an increasingly non-linear relation between slowness distribution and echo phase shift due to deviations from straight-ray propagation, but also by beamforming errors due to caustics (i.e. US propagating towards and from the same reflectors along multiple intersecting paths). We foresee that refraction can be compensated for, e.g. employing a more accurate forward model of sound propagation. Note that the key features of the new model, i.e. the CMA approach and the inverse-cosine law describing the relation between aberration delay and echo phase shift, are independent of the straight-ray approximation of how aberration delay is related to SoS (the latter was only proposed to enable real-time imaging). We therefore envisage that the new model of how aberration delays are reflected in echo phase shift is more generally applicable, e.g. with a bent-ray model of sound propagation. The influence of refraction on the final SoS image as well as potential remedies is an interesting topic, it is, however, beyond the scope of this study.

Another fundamental assumption was that of sharply defined Tx/Rx angles, but finite angular apertures are needed and were used in the experiments to enable a laterally resolved echo phase shift measurement (and thus SoS image). One may argue that the round-trip times of echoes thus correspond to average slowness integrals over wedge-shaped areas rather than along thin lines. Wave field simulations (not shown) indicate, however, that the angular aperture radius used for Tx/Rx focusing (2.5°) together with the wavelength (0.2 mm at 5 MHz) result in a collimated beam rather than a focused one, with a roughly constant beam diameter of 2.5 mm. Within this resolution, the rays can therefore be approximated as lines. The good accuracy of our results demonstrates that the simplification to line integrals had only minor influence. The influence of Tx and Rx angular apertures on the SoS results is an interesting topic, but was beyond the scope of this study.

The focus of the present study was to validate the new model. This was done in phantom experiments, and – so far – we did not attempt to investigate the clinical accuracy of CUTE. As a preliminary confirmation of the phantom results in an *in − vivo* case, however, Fig. 14 shows the B-mode and SoS images reconstructed with the old and the new model, of a healthy volunteer's liver (male, age 40, in compliance with the ethical principles of the Declaration of Helsinki 2018). As opposed to the old model, the new model reconstructs a spatial distribution of SoS that correctly delineates the spatial distribution of the different tissue types, and shows the expected uniform distribution of SoS inside

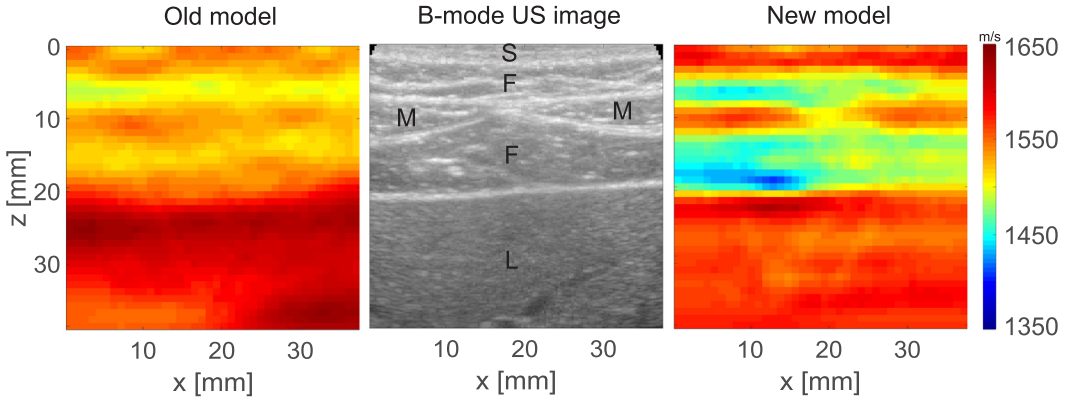

Fig. 14. Side-by-side display of conventional B-mode US of a healthy volunteer's abdominal wall with subjacent liver tissue (S: skin, F: fat, M: rectus abdominis muscle, L: liver parenchyma) and SoS images reconstructed with the old and the new model. In the new model, the SoS values of different tissues are well distinguished.

the liver. Moreover, the reconstructed mean SoS of the liver (1564±4 m/s) is reasonable for healthy liver tissue [52], making the new model a promising step towards quantitative hepatic imaging. Imbault et al. [25] reported that the average SoS of the liver of healthy patients is ≈ 1570 m/s. For patients having steatosis of grade 1 (according to the Brunt scale), the average SoS decreases by about 40 m/s. Thus, CUTE (with the proposed new model) has potential for a quantitative diagnosis of fatty liver disease.

To allow real time SoS imaging, the regularized pseudo-inverse matrix according to Eq. 21 can be pre-calculated and stored in a lookup table as described in the Materials and Methods section. As an example: applying the pre-calculated pseudo-inverse to the phase shift data takes only 0.04 s when implemented in Matlab® on the system with Intel Core I7-4790 CPU and 16 GB RAM. The bottle neck is the data transfer and beamforming (0.6 s when implemented with the Verasonics Vantage 64 LE and on the same PC, using the Verasonics DAS software) whereas the tracking takes 0.2 s on the same PC in Matlab®, totalling 0.8 s. Progress in hardware development (e.g. implementation on GPU) can thus easily allow several SoS images per second frame rate, making CUTE promising as an addition to real-time handheld US.

In summary, the new model is an important step towards quantitative clinical handheld reflection-mode SoS imaging. Thanks to the fast angle scanning of US systems and the low computational cost of CUTE, quantitative SoS images can be routinely displayed in parallel to conventional B-mode US. CUTE therefore can be used to image the SoS in any organ that can be examined using echo US. Furthermore, a more accurate estimation of the spatial distribution of SoS using the proposed new model will benefit an improved correction of conventional B-mode US images for SoS-related aberrations [38,39]. Apart from hepatic imaging for diagnosing steatosis, another promising target is the female breast with the goal to improve breast cancer diagnostics. Whereas using a linear probe is suitable for e.g. breast imaging, transabdominal hepatic imaging is conventionally performed using curvilinear probes operating at lower frequencies. The presented methodology can readily be adapted to such probes, however, the disadvantage of a curvilinear probe compared to a linear one is the smaller angle range that can be used for Tx/Rx beamsteering, as the same grid node is within the acceptance angle of a smaller number of elements when these elements are pointing into diverging directions. Also, the ratio between the probe aperture length and the required depth range for full liver imaging limits the available angle range, resulting in reduced axial resolution in comparison to superficial tissue. In situations where one is interested in a local average quantitative reading rather than a spatially resolved distribution of SoS, the low axial resolution may not be important. In situations where the lack of spatial resolution becomes a problem, using a very large aperture linear array probe instead of a curved one could be a viable solution for hepatic imaging.

CUTE is based on phase tracking and thus sensitive to motion artifacts. Our in vivo experience though shows that robust SoS images can be easily obtained when following a conventional breath hold procedure. When imaging near or inside a large blood vessel, however, blood flow and tissue clutter inhibit phase tracking and may cause SoS artefacts in the reconstruction. To overcome this limitation, we have already proposed a technique [53] where the first order echoes from the moving blood cells are separated from the static echo clutter via a clutter wall filter as in Doppler flow imaging. This technique extends the application of CUTE also to arteries, e.g. for the assessment of plaque inside the carotid artery and may also further improve liver imaging, since it can reduce SoS artefacts which may occur around hepatic veins.

Data availability

The experimental data used in this study are available on request from the corresponding Author.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research was funded in part by the Swiss National Science Foundation (project number 205320_178038) and the European Union's Horizon 2020 research and innovation programme under grant agreement No. 732411, Photonics Private Public Partnership, and is supported by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 16.0162. The opinions expressed and arguments employed herein do not necessarily reflect the official view of the Swiss Government. The authors thank René Nyffenegger for his technical assistance.

References

[1] J. Baker, P.J. Kornguth, M.S. Soo, R. Walsh, P. Mengoni, Sonography of solid breast lesions: observer variability of lesion description and assessment, AJR. Am. J. Roentgenol. 172 (1999) 1621–1625.
[2] K. Konno, H. Ishida, M. Sato, T. Komatsuda, J. Ishida, H. Naganuma, Y. Hamashima, S. Watanabe, Liver tumors in fatty liver: difficulty in ultrasonographic interpretation, Abdominal Imaging 26 (2001) 487–491.
[3] G. Rahbar, A.C. Sie, G.C. Hansen, J.S. Prince, M.L. Melany, H.E. Reynolds, V.P. Jackson, J.W. Sayre, L.W. Bassett, Benign versus malignant solid breast masses: Us differentiation, Radiology 213 (1999) 889–894.
[4] A. Athanasiou, A. Tardivon, M. Tanter, B. Sigal-Zafrani, J. Bercoff, T. Deffieux, J.-L. Gennisson, M. Fink, S. Neuenschwander, Breast lesions: quantitative elastography with supersonic shear imaging—preliminary results, Radiology 256 (2010) 297–303.
[5] J. Bamber, D. Cosgrove, C. Dietrich, J. Fromageau, J. Bojunga, F. Calliada, V. Cantisani, J.-M. Correas, M. D'onofrio, E. Drakonaki, et al., Efsumb guidelines and recommendations on the clinical use of ultrasound elastography. Part 1: Basic principles and technology, Ultraschall in der Medizin-Eur. J. Ultrasound 34 (2013) 169–184.
[6] R.G. Barr, Real-time ultrasound elasticity of the breast: initial clinical results, Ultrasound Quart. 26 (2010) 61–66.
[7] D. Cosgrove, F. Piscaglia, J. Bamber, J. Bojunga, J.-M. Correas, O. Gilja, A. Klauser, I. Sporea, F. Calliada, V. Cantisani, et al., Efsumb guidelines and recommendations on the clinical use of ultrasound elastography. part 2: Clinical applications, Ultraschall in der Medizin-Eur. J. Ultrasound 34 (2013) 238–253.
[8] C.F. Dietrich, J. Bamber, A. Berzigotti, S. Bota, V. Cantisani, L. Castera, D. Cosgrove, G. Ferraioli, M. Friedrich-Rust, O.H. Gilja, et al., Efsumb guidelines and recommendations on the clinical use of liver ultrasound elastography, update 2017 (long version), Ultraschall in der Medizin-Eur. J. Ultrasound 38 (2017) (2017) e16–e47.
[9] R.M. Sigrist, J. Liau, A. El Kaffas, M.C. Chammas, J.K. Willmann, Ultrasound elastography: review of techniques and clinical applications, Theranostics 7 (2017) 1303.
[10] S. Hu, L.V. Wang, Photoacoustic imaging and characterization of the microvasculature, J. Biomed. Opt. 15 (2010) 011101.
[11] M. Jaeger, D.C. Harris-Birtill, A.G. Gertsch-Grover, E. O'Flynn, J.C. Bamber, Deformation-compensated averaging for clutter reduction in epiphotoacoustic imaging in vivo, J. Biomed. Opt. 17 (2012) 066007.
[12] K.G. Held, M. Jaeger, J. Rička, M. Frenz, H.G. Akarçay, Multiple irradiation sensing of the optical effective attenuation coefficient for spectral correction in handheld oa imaging, Photoacoustics 4 (2016) 70–80.
[13] L. Ulrich, L. Ahnen, H. Akarçay, S. Sanchez, M. Jaeger, G. Held, M. Wolf, M. Frenz, Spectral correction for handheld optoacoustic imaging by means of near-infrared optical tomography in reflection mode (2018).
[14] N. Ruiter, M. Zapf, R. Dapp, T. Hopp, W. Kaiser, H. Gemmeke, First results of a clinical study with 3d ultrasound computer tomography, in: 2013 IEEE International Ultrasonics Symposium (IUS), IEEE, 2013, pp. 651–654.
[15] J.F. Greenleaf, R.C. Bahn, Clinical imaging with transmissive ultrasonic computerized tomography, IEEE Trans. Biomed. Eng. (1981) 177–185.
[16] J. Jago, T. Whittingham, Experimental studies in transmission ultrasound computed tomography, Phys. Med. Biol. 36 (1991) 1515.
[17] G. Zografos, D. Koulocheri, P. Liakou, M. Sofras, S. Hadjiagapis, M. Orme, V. Marmarelis, Novel technology of multimodal ultrasound tomography detects breast lesions, Eur. Radiol. 23 (2013) 673–683.
[18] P.L. Carson, C.R. Meyer, A.L. Scherzinger, T.V. Oughton, Breast imaging in coronal planes with simultaneous pulse echo and transmission ultrasound, Science 214 (1981) 1141–1143.
[19] P. Huthwaite, F. Simonetti, High-resolution imaging without iteration: A fast and robust method for breast ultrasound tomography, J. Acoust. Soc. Am. 130 (2011) 1721–1734.

[20] J. Wiskin, D. Borup, S. Johnson, M. Berggren, Non-linear inverse scattering: High resolution quantitative breast tissue tomography, J. Acoust. Soc. Am. 131 (2012) 3802–3813.
[21] G. Sandhu, C. Li, O. Roy, S. Schmidt, N. Duric, Frequency domain ultrasound waveform tomography: breast imaging using a ring transducer, Phys. Med. Biol. 60 (2015) 5381.
[22] M.C. Hesse, L. Salehi, G. Schmitz, Nonlinear simultaneous reconstruction of inhomogeneous compressibility and mass density distributions in unidirectional pulse-echo ultrasound imaging, Phys. Med. Biol. 58 (2013) 6163.
[23] H.-C. Shin, R. Prager, H. Gomersall, N. Kingsbury, G. Treece, A. Gee, Estimation of average speed of sound using deconvolution of medical ultrasound data, Ultrasound Med. Biol. 36 (2010) 623–636.
[24] J. Krucker, J.B. Fowlkes, P.L. Carson, Sound speed estimation using automatic ultrasound image registration, IEEE Trans. Ultrasonics Ferroelectr. Frequency Control 51 (2004) 1095–1106.
[25] M. Imbault, A. Faccinetto, B.-F. Osmanski, A. Tissier, T. Deffieux, J.-L. Gennisson, V. Vilgrain, M. Tanter, Robust sound speed estimation for ultrasound-based hepatic steatosis assessment, Phys. Med. Biol. 62 (2017) 3582.
[26] M. Kondo, K. Takamizawa, M. Hirama, K. Okazaki, K. Iinuma, Y. Takehara, An evaluation of an in vivo local sound speed estimation technique by the crossed beam method, Ultrasound Med. Biol. 16 (1990) 65–72.
[27] I. Céspedes, J. Ophir, Y. Huang, On the feasibility of pulse-echo speed of sound estimation in small regions: Simulation studies, Ultrasound Med. Biol. 18 (1992) 283–291.
[28] M. Jakovljevic, S. Hsieh, R. Ali, G. Chau Loo Kung, D. Hyun, J.J. Dahl, Local speed of sound estimation in tissue using pulse-echo ultrasound: Model-based approach, J. Acoust. Soc. Am. 144 (2018) 254–266.
[29] M. Jaeger, G. Held, S. Peeters, S. Preisser, M. Grünig, M. Frenz, Computed ultrasound tomography in echo mode for imaging speed of sound using pulse-echo sonography: proof of principle, Ultrasound Med. Biol. 41 (2015) 235–250.
[30] M. Jaeger, M. Frenz, Towards clinical computed ultrasound tomography in echo-mode: Dynamic range artefact reduction, Ultrasonics 62 (2015) 299–304.
[31] T. Loupas, J. Powers, R.W. Gill, An axial velocity estimator for ultrasound blood flow imaging, based on a full evaluation of the doppler equation by means of a two-dimensional autocorrelation approach, IEEE Trans. Ultrasonics Ferroelectrics Frequency Control 42 (1995) 672–688.
[32] S. Flax, M. O'Donnell, Phase-aberration correction using signals from point reflectors and diffuse scatterers: Basic principles, IEEE Trans. Ultrasonics Ferroelectrics Frequency Control 35 (1988) 758–767.
[33] L. Nock, G.E. Trahey, S.W. Smith, Phase aberration correction in medical ultrasound using speckle brightness as a quality factor, J. Acoust. Soc. Am. 85 (1989) 1819–1833.
[34] D. Rachlin, Direct estimation of aberrating delays in pulse-echo imaging systems, J. Acoust. Soc. Am. 88 (1990) 191–198.
[35] Y. Li, Phase aberration correction using near-field signal redundancy. i. principles [ultrasound medical imaging], IEEE Trans. Ultrasonics Ferroelectrics Frequency Control 44 (1997) 355–371.
[36] Y. Li, D. Robinson, D. Carpenter, Phase aberration correction using near-field signal redundancy. II. Experimental results, IEEE Trans. Ultrasonics Ferroelectrics Frequency Control 44 (1997) 372–379.
[37] M.A. Haun, D.L. Jones, W. Oz, Overdetermined least-squares aberration estimates using common-midpoint signals, IEEE Trans. Med. Imaging 23 (2004) 1205–1220.
[38] M. Jaeger, E. Robinson, H.G. Akarçay, M. Frenz, Full correction for spatially distributed speed-of-sound in echo ultrasound based on measuring aberration delays via transmit beam steering, Phys. Med. Biol. 60 (2015) 4497.
[39] H.-M. Schwab, A. Ihrig, D. Depke, S. Hermann, M. Schäfers, G. Schmitz, Aberration correction in photoacoustic imaging using paraxial backpropagation, in: 2017 IEEE International Ultrasonics Symposium (IUS), IEEE, 2017, pp. 1–4.
[40] M. Jaeger, M. Frenz, Quantitative imaging of speed of sound in echo ultrasonography, in: iEEE International Ultrasound Symposium, Taipei, 2015. https://www.youtube.com/watch?v=Ck75XbfLQtY.
[41] S.J. Sanabria, E. Ozkan, M. Rominger, O. Goksel, Spatial domain reconstruction for imaging speed-of-sound with pulse-echo ultrasound: simulation and in vivo study, Phys. Med. Biol. 63 (2018) 215015.
[42] L. Ruby, S.J. Sanabria, K. Martini, K.J. Dedes, D. Vorburger, E. Oezkan, T. Frauenfelder, O. Goksel, M.B. Rominger, Breast cancer assessment with pulse-echo speed of sound ultrasound from intrinsic tissue reflections: Proof-of-concept, Investigative Radiol. 54 (2019) 419–427.
[43] J.B. Bednar, Modeling, migration and velocity analysis in simple and complex structure, by Panorama Technologies, Inc, 2009.
[44] R. Rau, D. Schweizer, V. Vishnevskiy, O. Goksel, Speed-of-sound imaging using diverging waves, arXiv preprint arXiv:1910.05935, 2019.
[45] W.F. Walker, G.E. Trahey, The application of k-space in pulse echo ultrasound, IEEE Trans. Ultrasonics Ferroelectrics Frequency Control 45 (1998) 541–558.
[46] J. Bercoff, Ultrafast ultrasound imaging, in: Ultrasound imaging-Medical applications, IntechOpen, 2011.
[47] G. Montaldo, M. Tanter, J. Bercoff, N. Benech, M. Fink, Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography, IEEE Trans. Ultrasonics Ferroelectrics Frequency Control 56 (2009) 489–506.
[48] E.L. Madsen, J.A. Zagzebski, G.R. Frank, An anthropomorphic ultrasound breast phantom containing intermediate-sized scatteres, Ultrasound Med. Biol. 8 (1982) 381–392.
[49] E. Madsen, G. Frank, T. Krouskop, T. Varghese, F. Kallel, J. Ophir, Tissue-mimicking oil-in-gelatin dispersions for use in heterogeneous elastography phantoms, Ultrasonic Imaging 25 (2003) 17–38.
[50] E.L. Madsen, M.A. Hobson, G.R. Frank, H. Shi, J. Jiang, T.J. Hall, T. Varghese, M.M. Doyley, J.B. Weaver, Anthropomorphic breast phantoms for testing elastography systems, Ultrasound Med. Biol. 32 (2006) 857–874.
[51] M.M. Nguyen, S. Zhou, J.-L. Robert, V. Shamdasani, H. Xie, Development of oil-in-gelatin phantoms for viscoelasticity measurement in ultrasound shear wave elastography, Ultrasound Med. Biol. 40 (2014) 168–176.
[52] T. Lin, J. Ophir, G. Potter, Correlations of sound speed with tissue constituents in normal and diffuse liver disease, Ultrasonic Imaging 9 (1987) 29–40.
[53] M. Kuriakose, J.-W. Muller, P. Stähli, M. Frenz, M. Jaeger, Receive beam-steering and clutter reduction for imaging the speed-of-sound inside the carotid artery, J. Imaging 4 (2018) 145.